Wind farm parameterization and turbulent wind box generation

Michael Mc William, Nicolas Bonfils, Nikolay Dimitrov, Suguang Dou

To cite this version:

Michael Mc William, Nicolas Bonfils, Nikolay Dimitrov, Suguang Dou. Wind farm parameterization and turbulent wind box generation. DTU; IFPEN. 2022, pp.Deliverable n° D3.1. hal-04033050

HAL Id: hal-04033050
https://ifp.hal.science/hal-04033050

Submitted on 16 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HIPERWIND

HIghly advanced Probabilistic design and Enhanced Reliability methods for high-value, cost-efficient offshore WIND

Title: Wind farm parameterization and turbulent wind box generation

Deliverable no: D3.1

Delivery date: 31.03.2022
Lead beneficiary: DTU
Dissemination level: Public

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 101006689

Author(s) information (alphabetical):

Name	Organisation	Email
Michael McWilliam	DTU	mimc@dtu.dk
Nicolas Bonfils	IFPEN	nicolas.bonfils@ifpen.fr
Nikolay Dimitrov	DTU	nkdi@dtu.dk
Suguang Dou	DTU	sudou@dtu.dk

Acknowledgements/Contributions:

Name	Name	Name
Mark Kelly, DTU		
Martin Guiton, IFPEN		
Ásta Hannesdóttir, DTU		

Document information:

Version	Date	Description	Prepared by	Reviewed by	Approved by
0	31-03-2022	First official version	Authors listed above	Mark Kelly, Martin Guiton	
1.0	19-01-2023	Minor corrections introduced			

Definitions:

Contents

1. Executive Summary 2
2. Introduction 5
  2.1. Wind farm wake effects parameterization 6
  2.2. Advanced turbulence box generation 9
3. Wind farm wake effect parameterization using layout encoding 12
  3.1. Proposed layout parameterization approaches 12
  3.2. Assessment of layout parameterization approaches using Pywake 14
  3.3. Load surrogates with the DWM model 18
  3.4. Comparison with SCADA measurements 28
4. Wind farm wake effect parameterization using surrogates for local wind estimates 33
  4.1. Introduction and objective 33
  4.2. Parametrization approach 35
    4.2.1. Description of the model function 35
    4.2.2. Wind field plane fitting methodology 38
  4.3. Design of experiments 43
    4.3.1. FarmShadowTM DoE 43
    4.3.2. Fitting process 45
    4.3.3. Design of experiments in terms of parameters 50
  4.4. DLW simulations 51
    4.4.1. Model description 51
    4.4.2. Wind modifications 51
    4.4.3. Turbulence spectra parameters 53
    4.4.4. Output computation (DEL and extreme loads) 53
  4.5. Surrogate model and GSA 54
    4.5.1. Kriging 54
    4.5.2. Kriging performance 55
    4.5.3. Global Sensitivity Analysis (GSA) 57
  4.6. Application case 58
  4.7. Conclusion and perspectives 64

5. Comparison between wake parameterization method outputs 66
6. Transient wind event simulations using a constrained turbulence approach 68
  6.1. The Hipersim/turbgen tool for turbulence box generation and constrained turbulence simulations 68
  6.2. Impact of transient event profiles on turbine loads 70
7. Non-Gaussian turbulence time series generation for load simulations 90
  7.1. Non-Gaussian turbulence box generation based on target spectra and statistical moments 90
  7.2. Impact of non-Gaussian turbulence boxes on turbine loads 94
8. Conclusions 98

List of Abbreviations

DEL	Damage-equivalent fatigue load
DoE	Design of Experiment
DWM	Dynamic Wake Meandering
ECD	Extreme Coherent gust with Direction change
EDC	Extreme Direction Change
EOG	Extreme Operating Gust
ETM	Extreme Turbulence Model
FEA	Finite Element Analysis
FFT	Fast Fourier Transform
FLS	Fatigue Limit State
FNN	Feedforward Neural Network
GSA	Global Sensitivity Analysis
IEC	International Electrotechnical Committee
IFFT	Inverse Fast Fourier Transform
LCOE	Levelized Cost of Energy
LES	Large Eddy Simulations
LHS	Latin Hypercube Sampling
MC	Monte Carlo
NRMSE	Normalized Root Mean Squared Error
NTM	Normal Turbulence Model
PCE	Polynomial Chaos Expansion
QoI	Quantity of Interest
RBF	Radial Basis Function
RMSE	Root Mean Squared Error
SCADA	Supervisory Control And Data Acquisition
std	Standard deviation
ULS	Ultimate Limit State

1. Executive Summary

The present report addresses Hiperwind Deliverable 3.1: wind farm parameterization and turbulent wind box generation. This includes two separate scopes: 1) studying the effect of wakes on wind turbine loads - and more specifically how to describe (parameterize) the wake effects in terms of quantities suitable for surrogate model inputs, and 2) turbulent wind field generation for special scenarios including transient events (extreme wind gusts, ramps and direction changes) and situations with non-Gaussian statistics of the wind field. The common for both topics is that they represent aspects of the inflow of wind turbines in wind farms that are a major factor affecting the wind turbine loads. Thus, the main purpose of this study is to develop methods and tools for more efficient and accurate modelling of such scenarios, and to better understand the uncertainties involved and the relative significance of the different inputs on wind turbine loads and power.

Summary: wind parameterization in a wind farm

We studied two wind farm parameterization methods that are suitable for two different workflows for wind farm load assessment. In the first method developed by DTU, wind turbine loads are simulated using the Hawc2 aeroelastic tool, with the Dynamic Wake Meandering (DWM) model used to include a dynamic simulation of incoming wakes. In order to define appropriate inputs for a surrogate model that can replace the Hawc2+DWM simulations, the wake effects are described in terms of variables related to the geometric layout of the wind farm such as relative wind turbine positions. Several parameterization approaches were tested within this workflow: encoding based on the overlap between the wake deficit and the rotor, on the relative upwind turbine positions, and on an autoencoder model that applies dimensionality reduction on the relative position encoding. The encoding obtained in this way served as an input to a regression model based on Feedforward Neural Networks.

The second parameterization method developed by IFPEN is based on using shape functions to describe the wake deficit properties. First, a "wake parameterization" surrogate model is trained that uses a few parameters to describe a mixture of Gaussian shape functions that define the averaged properties of the wake deficit under a broad range of waked inflow conditions. This surrogate is trained on simulations with the FarmShadow tool. Then, the predicted wake deficit is superimposed on a turbulent wind field and fed into load simulations with Deeplines Wind. Finally, a second surrogate model is trained that maps the wake parameterization surrogate outputs to the outputs of load simulations. Combining the two models, the loads can be predicted for a wind turbine in an arbitrary wind farm layout.

Sensitivity analysis in terms of Sobol indices was carried out with both parameterization approaches, and the results from the load assessment and sensitivity analysis were compared. The outcome of the wake parameterization studies led to the following conclusions:

- It was shown that it is feasible to parameterize the wind farm layout in a way that enables the construction of surrogate models that predict loads and power outputs of individual turbines in arbitrary wind farm layouts;
- The sensitivity analysis showed that the ambient wind speed is the governing factor for loads and power production. The wind direction (which determines the strength of the wake effects) and the ambient turbulence are also of high importance. For wind speeds just below rated where the turbine thrust is highest, the wind direction has higher importance than the turbulence, while at low wind speeds the turbulence has more significant effect. For all methods that were studied, the wind shear was found to be of small significance.

- The performance of the surrogates is affected by the significance of wake effects: for signals with less wake impact, such as the extremes, the performance is consistently higher

Summary: uncertainty reduction in turbulent wind box generation

The second inflow aspect we considered is the turbulent field that is used as input to load simulations. The focus is on situations that deviate from the normal operation under stationary conditions where the standard turbulence generation approaches provide sufficient quality of results. The special scenarios considered were transient events (extreme wind gusts, ramps and direction changes), and situations with highly non-Gaussian field statistics. For the purpose of this investigation, an open Python-based turbulence field generation tool was created within the Hipersim software package. The tool generates turbulence fields that can serve as inputs to aeroelastic load simulations, and can embed measured wind time series as constraints in the wind field. The capabilities were further expanded by introducing the possibility to generate non-Gaussian fields with a predefined skewness and kurtosis.

Using the constrained simulation approach, the load response of the DTU10MW turbine was simulated under a set of transient events that were previously obtained from measurements as part of Hiperwind Deliverable 2.3. The load results from these events were compared to simulations with the transient design load cases defined in the IEC61400-1 standard. In addition, the effect of introducing non-Gaussian statistics was assessed by carrying out pairs of load simulations with the same turbulence seeds but different skewness and kurtosis. The following conclusions were drawn:

- It was verified that the Hipersim/Turbgen tool produces turbulence boxes with the correct spectrum and coherence properties as prescribed with the Mann turbulence spectrum;
- The constrained simulation functionality of Hipersim results in correct reproduction of the target time series in the turbulence boxes;
- Overall, the loads obtained using simulations with measured transient events were in most cases less than the loads resulting from the synthetic events prescribed by the IEC standards. However, there are also cases where the constrained turbulence loads were either similar or exceeded the IEC 61400 loads.
- The synthetic conditions from the IEC standard tend to produce a different relationship between wind speed and load magnitude, compared to the measured time series simulations. The highest loads from IEC events tended to be concentrated near rated wind conditions, while the highest constrained turbulence loads were more widely distributed over more wind speeds.
- A non-Gaussian turbulence box generator was developed that produces turbulence boxes with the Veers model and arbitrary skewness and kurtosis. It was verified that the resulting turbulence fields attain the target statistical moments, while retaining the correct spectrum and coherence.
- The present method of non-Gaussian turbulence generation is based on statistical rather than physical considerations. Although the results are statistically consistent, there is a chance that they are unphysical as e.g. major intermittent wind direction reversals. Increasing the skewness and kurtosis of a wind time series amplifies the chance of such non-physical events.
- For most load channels, the effect of non-Gaussian wind statistics was mild, such as minor changes in the design driving load and a shift in the wind speed where the dominant load events occur. An exception were the tower top yaw moment extremes, where changes in the skewness and kurtosis led to a significant increase in the loads.

Research significance

The research activities within this study provide solutions for more accurate and more computationally efficient assessment of loads under challenging scenarios such as wakes, transients, and unusual turbulence structures that have non-Gaussian statistics. This provides the tools for more efficient wind turbine and wind farm design process, and opens possibilities for uncertainty-aware design and optimization.

In addition, the outcomes of this work increase our understanding of the significance of wind turbine wakes, ambient wind conditions and transients on wind turbine loads. This provides insights into what type of scenarios need increased focus in future design and planning activities, and potentially will influence the future versions of wind turbine design standards.

2. Introduction

The HIPERWIND project (HIghly advanced Probabilistic design and Enhanced Reliability methods for high-value, cost efficient offshore WIND) aims at significantly reducing the levelized cost of energy (LCOE) of offshore wind farms, through advancements of basic wind energy science which will lead to reductions in risk and uncertainty. The outcome is cost efficient offshore wind through a reduction in unnecessary use of materials, less unscheduled maintenance, and optimized operating strategy tailored at delivering power with high market value. The project concept is outlined in Figure 1.

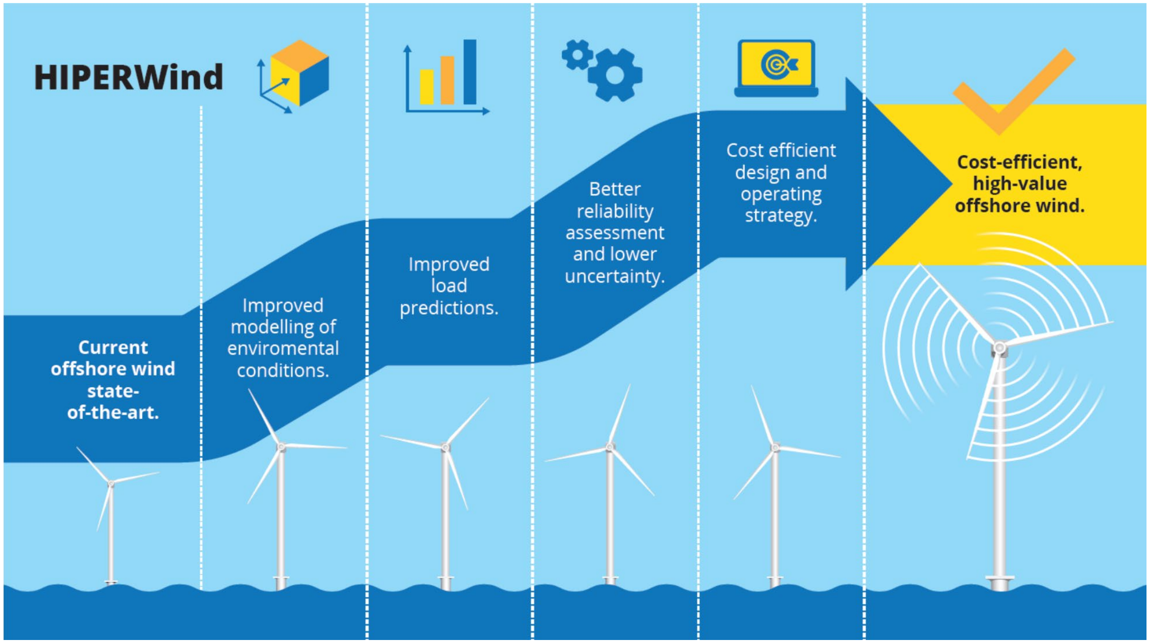


Figure 1 Hiperwind concept

The main convergence point of all project activities is the goal of making probabilistic design feasible. This happens through uncertainty quantification and reduction, improving the efficiency and accuracy of the models in the design chain, and introducing uncertainty-aware methods and tools. This improvement drive happens on several different scales: starting from the atmospheric conditions, through local conditions on wind farm level and their effect on wind turbine loads, and finally focusing on the wind turbine where the main scope is structural and component reliability and how these are affected by the environmental conditions and wind farm interactions.

Work package 3 in Hiperwind deals with characterizing the uncertainties in load predictions for offshore wind turbines in a wind farm. A significant share of these uncertainties arise from the turbulent wind field and the way it is modified by the wakes of upwind turbines. The wind field and its effect on loads is the primary scope of Hiperwind Deliverable 3.1 which is addressed with this report.

The main purpose of the study is to provide solutions for more accurate and more computationally efficient assessment of loads under challenging scenarios such as wakes, transients, and unusual turbulence structures that have non-Gaussian statistics. In addition, we would like to better understand the uncertainties involved and the relative significance of the different inputs on wind turbine loads and power. This is addressed through the following specific research goals:

- Develop and validate methods for wind farm wake effect parameterization, to enable the construction of efficient surrogate models that can provide accurate load and power predictions for individual turbines in an arbitrary wind farm layout, with using just ambient wind conditions and the wind farm layout as inputs. Carry out sensitivity analysis to outline the relative importance of the model inputs on wind turbine design loads.
- Develop and test tools capable of simulating the load effects on wind turbines in scenarios with challenging wind conditions such as atmospheric transients and non-Gaussian wind field properties. Demonstrate the load simulation procedures with such scenarios and compare the load outcomes with those from standard IEC design load cases.

2.1. Wind farm wake effects parameterization

Flow disturbances from wind turbine wakes have a significant effect on the loads and power production in wind farms, potentially leading to energy losses in the order of 10% and a reduction in fatigue lifetime [1]. Within wind energy, the scientific community and industry are continuously working on understanding and controlling wake effects through developing improved wake models and simulation tools as well as focusing on layout and control strategy optimization [2]. The flow in the wind turbine wake is characterized with a wind speed reduction (deficit) compared to the ambient flow, and with changes in the turbulence (increased variance and modified scales and shapes of the turbulence structures). The wake characteristics (such as the amount, shape, and extent of the wind speed reduction) are strongly dependent on the rotor induction and thrust of the source turbine, which is in turns dependent on the ambient conditions. Further, the wake disturbances evolve downwind from the source wind turbine, meaning that their characteristics will be dependent on how the ambient conditions affect the turbulent mixing of the wakes with the surrounding fluid, but also be a function of the relative position between the wake source and the affected turbine. As a result, in large wind farms (and especially those with irregular layout) there can be a large number of unique wake situations, depending on the combinations of ambient wind conditions, wind direction and turbine positions. Properly characterizing the wind farm power output and loading conditions over the full operational envelope requires that all these unique wake situations are taken into account.

The typical workflow for wind farm power and load assessment is illustrated in Figure 2. Mapping the power output can be done in a computationally efficient way, using quasi-static wake models that predict the average local wind conditions at the rotor of each individual turbine - and subsequently the average power output. The loads assessment however will typically require time-domain simulations that are significantly more expensive computationally. This makes it impractical to sample the full range of wake situations for all wind turbines. As an example, for a medium-size wind farm with 30 turbines, sampling the wind direction with 2deg resolution, and the wind speed with 1 m/s resolution from 4 to 25 m/s, will result in a total of 30*180*22 = 118800 different scenarios. Given that it takes at least 1 h of simulations to reduce the uncertainty due to realization-to-realization variations in the turbulence fields used as simulation inputs, we also need 118800 hours of simulation to complete the sampling. Since a typical load simulation tool runs at approximately 1:1 speed compared to real time, the entire sample will require the same amount of CPU hours as the hours of simulation outputs.

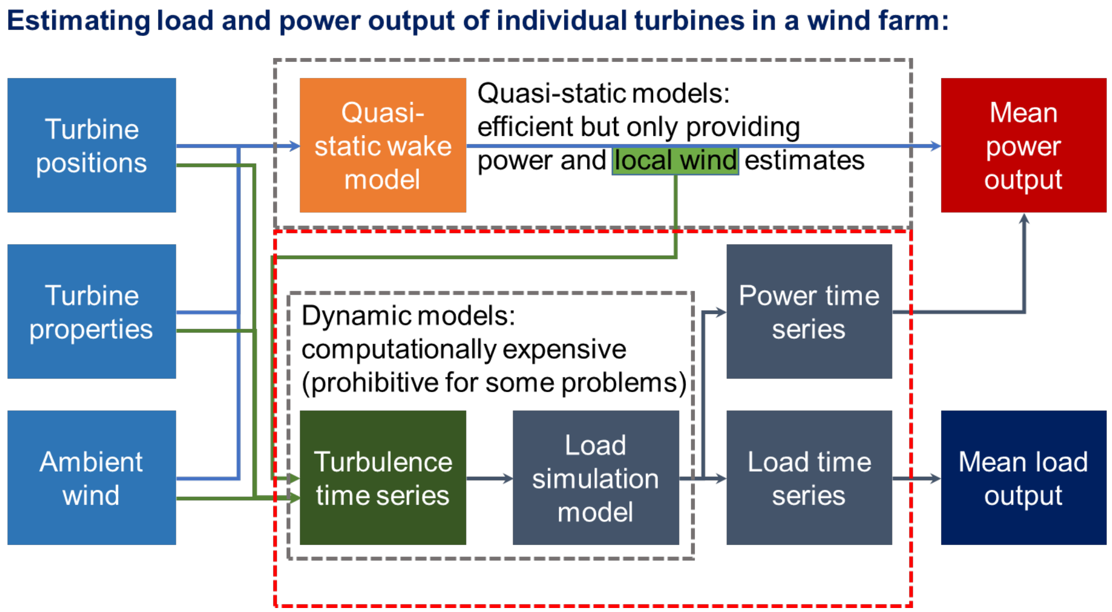


Figure 2 Diagram showing the power and load statistics estimation process for wind turbines in a wind farm.

A popular approach for reducing the computational expenses is training a surrogate model that approximates the behavior of the complex aeroelastic model but has much better computational efficiency [3] – effectively replacing the dynamic system enclosed by the red dashed line in Figure 2 with a functional approximation. As a general definition[1], for a pair of inputs $\mathbf{x}$ and outputs $\mathbf{y}$, a surrogate model $g(\mathbf{x}, \boldsymbol{\theta})$ represents the relationship

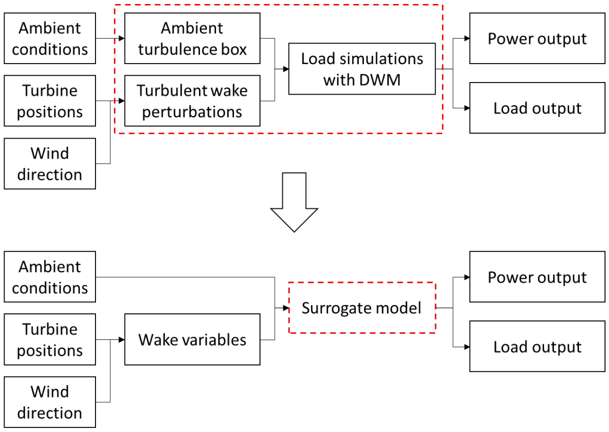
$$\mathbf{y} + \boldsymbol{\varepsilon}_y = g(\mathbf{x}, \boldsymbol{\theta}) + \boldsymbol{\varepsilon}_g$$

where $\boldsymbol{\theta}$ are the model parameters, $\boldsymbol{\varepsilon}_g$ is a model error accounting for the limitations in the model function which lead to imperfect mapping between inputs and outputs, and $\boldsymbol{\varepsilon}_y$ is a variation (error) in the outputs due to factors external to the surrogate model, such as e.g. measurement uncertainty or realization-to-realization uncertainty from the input turbulence fields. The model is trained by determining the values of $\boldsymbol{\theta}$ that provide the best accuracy, typically by minimizing the variance of the error $\boldsymbol{\varepsilon}_g$. The trained surrogate model represents a unique mapping between $\mathbf{x}$ and $\mathbf{y}$. Obviously, when dealing with wind farms, the wake effects need to be integrated in the surrogate model inputs in terms of a fixed set of quantities that is a subset of $\mathbf{x}$. This is partially achieved by including variables descriptive of the ambient conditions such as wind speed, turbulence intensity, and wind shear. The other major factor governing the wake effects is the wind farm layout. If we want the surrogate model to be generalizable (i.e., applicable to more than one wind farm), the wind farm layout needs to be **parameterized**, i.e., described by a set of derived quantities that are not unique to a given wind farm. The parameterization can target the most accurate representation of either the wake-affected wind field (such as deficit depth and shape) or directly the wind turbine load and power outputs. There are several suggested parameterization approaches in literature (such as [4] and [5]), however there are no studies that systematically assess and compare such parameterization approaches. In Sections 3 and 4 of this report, we address the wake parameterization problem, focusing on two partially different approaches associated with two representative toolchains:

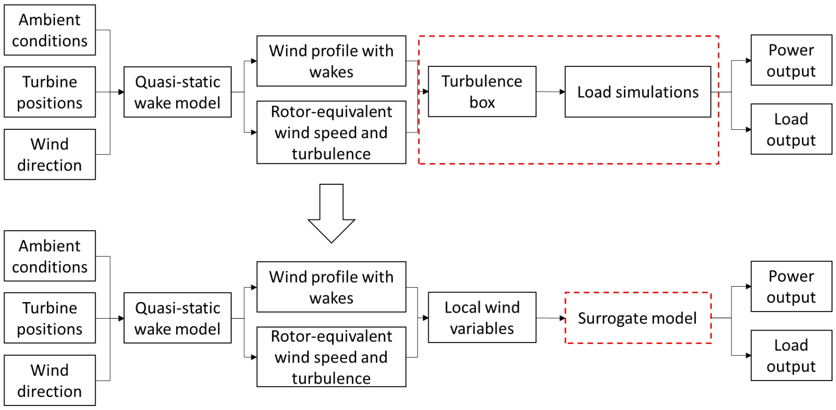
---

[1] In the above, and throughout the document, we use bold symptoms such as $\mathbf{x}$ to indicate that the variables represent multi-dimensional vectors.

1) Wind farm wake effect parameterization based on layout encoding. With this approach, the wake-related variables encode the relative positions of wake sources with respect to the wind turbine under consideration. This is convenient as an input to load simulations within the DTU toolchain based on the Dynamic Wake Meandering (DWM) model [6], where each upwind turbine is considered as a wake source and its effect is included in the turbulence input to the simulations, based on the relative location of the source and the propagation of its effect downwind. Therefore, when using the wake variables based on layout encoding, the surrogate model can replace the part of the toolchain involved in turbulence generation, wake meandering simulation, and load simulation (see sketch below). This approach is considered in Section 3 of this report.



2) Wind farm wake effect parameterization using surrogates for local wind estimates. This approach addresses a different toolchain used by IFPEN where the wake sources are not explicitly included in the dynamic load simulations (see sketch below). Instead, a quasi-static wake model is used to estimate a representative wind profile that includes the averaged wake effects. This profile is then incorporated in a single turbulence box which is fed to the load simulation tool. With this modelling approach, the step from wind farm layout to local wind profile is very computationally efficient and does not need to be replaced by a surrogate. Instead, a surrogate model for load simulations uses a description of the local wind conditions as an input, to predict the loads and power output. Hence, with this toolchain the main focus is to determine what are the variables that provide the best description of the local wind conditions including averaged wake effects. This approach is considered in Section 4 of this report.



The studies presented in Sections 3 and 4 use the Teesside offshore wind farm (Figure 3) as a study case. The primary focus is on Turbine #14 (shown with red dot in Figure 3) because it is located in the center of the wind farm and is subject to complex wake disturbances.

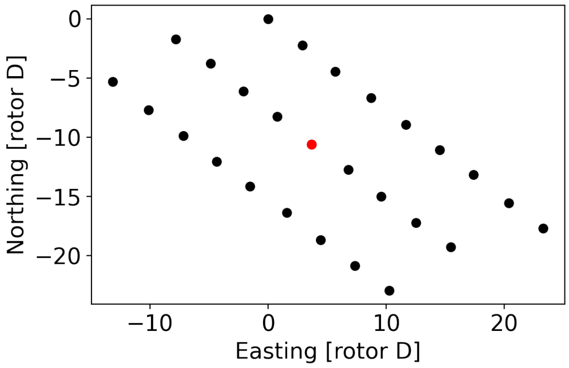


Figure 3 Teesside wind farm layout illustration. The red dot indicates turbine #14 which is used as reference in the wake effect studies in this report. A more detailed map is available at https://www.offshorewind.biz/2013/01/17/uk-searoc-appointed-principal-contractor-on-teesside-offshore-wind-farm/

Having computationally efficient surrogate models that can simultaneously capture the effect of ambient conditions and wakes provides additional opportunities. Namely, variance-based sensitivity analysis (analyzing the contribution of each input variable to the variance of the output) can be carried out, to evaluate the significance of wake effects and compare their impact to that of environmental variables such as wind speed and turbulence intensity. For some specific surrogate modelling approaches such as Polynomical Chaos Expansions (PCE), it is possible to directly derive the contribution of each input to the output variance based on the model coefficients [3],[7]. For some other approaches such as Feed-forward Neural Networks (FNNs), a first-order relationship between inputs and outputs can be computed analytically from the model coefficients [4]. The common property of most surrogate models is the computational efficiency, which allows intensive sampling from the models. It is therefore convenient to carry out variance-based sensitivity analysis by running Monte Carlo simulations on the surrogate models using the approach described in [8]. We apply the latter approach to obtain Sobol indices as measures of global variance-based sensitivity, with results reported in Section 3 and 4.

2.2. Advanced turbulence box generation

The current wind turbine design practice relies on load simulations carried out according to design scenarios recommended by standards such as the IEC 61400-1[9]. The large number of design simulations requires that the computations are done in an efficient way. The wind fields are generated with engineering models, where efficiency comes with the tradeoff of making some simplifying assumptions about the properties of the wind field. The standard tools such as Turbsim [10] and the Mann turbulence generator [11] built into Hawc2 [12] provide stationary, Gaussian wind fields, with uniform variance and predefined spectrum and coherence properties.

While the Normal (Gaussian) probability distribution has useful properties that allows easier mathematical formulation and development of such generative models, the point statistics of a turbulent wind field are not necessarily Gaussian. For normal operation scenarios under nominal wind conditions, the Gaussian assumption is found to result in reasonable loading as seen in [13]. However, the non-Gaussian properties are especially apparent in special situations such as the transient events shown in Deliverable 2.3 in Hiperwind [14], extreme turbulence events, and similar. Non-Gaussian atmospheric turbulence fields can be simulated with Large Eddy Simulations (LES), however this comes with significant computational expense. The possibility to simulate non-Gaussian fields in a computationally efficient way is therefore an important research aspect [15]. In the present report, the focus is on developing computationally efficient methods to generate non-Gaussian turbulence fields, based on the concept of generating a Gaussian turbulence box using standard engineering approaches, which is then subjected to mathematical transformations to produce a non-Gaussian field.

Another uncertain aspect of the wind field definitions in the IEC design load scenarios are the formulations of synthetic transient events that are used for ultimate limit state (ULS) design load assessment. The Extreme Operating Gust (EOG), Extreme Direction Change (EDC) and Extreme Coherent gust with Direction change (ECD) events are based on analytical formulas describing the changes in wind field properties. These event profiles do not necessarily match the actual transient events in the atmosphere, and there are indications that this may result in unrealistic load assessments [16],[17]. In Deliverable 2.3 of the Hiperwind project, we developed a procedure for detecting and isolating significant transient events, and carried out analysis on 10+ years of data from the Høvsøre test station in Denmark. In the present study, we take the next step in analyzing what is the significance of such transient events on wind turbine loads, and how does the load response compare to that from simulations with the synthetic transient events defined by the IEC standards. The transient events extracted from the Høvsøre data are included as time series in turbulence boxes using the constrained turbulence simulation approach as defined in [18]. These turbulence boxes are then fed as inputs to aeroelastic load simulations, and the resulting loads are compared to those from simulations with the synthetic IEC events.

Implementing the methods for advanced turbulence field generation requires significant modifications to the programming scripts in existing turbulence generation solutions. In order to have full flexibility in this process, all development is done within a newly created turbulence generation package in Python called Turbgen, which is a submodule of the Hipersim package [19]. First, the standard Mann model generation based on the algorithm in Mann (1998) is implemented, followed by modifications for adding constrained simulation capabilities and non-Gaussian turbulence generation capabilities.

Several sections later in this report describe the details of the advanced turbulence box generation studies:

- Section 6 deals with transient events simulations with the constrained turbulence approach, and more specifically
    - Section 6.1 presents the Hipersim/Turbgen turbulence box generation tool and its capabilities of including transient event time series through constrained simulation;
    - Section 6.2 presents an analysis of the effect of transient events on loads;
- Section 7 describes the methods for non-Gaussian turbulence box generation and shows a brief analysis of how the changes in skewness and kurtosis affect the wind turbine load results.

References in Section 2

[1] R.J. Barthelmie, K. Hansen, S.T. Frandsen, O. Rathmann, J.G. Schepers, W. Schlez, et al. (2009) Modelling and measuring flow and wind turbine wakes in large wind farms offshore. Wind Energy 12, pp. 431-444

[2] van Wingerden, J. W., Fleming, P. A., Göçmen, T., Eguinoa, I., Doekemeijer, B. M., Dykes, K., … Giebel, G. (2020). Expert Elicitation on Wind Farm Control. Journal of Physics: Conference Series, 1618(2), 022025. https://doi.org/10.1088/1742-6596/1618/2/022025

[3] Dimitrov, N. K., Kelly, M. C., Vignaroli, A., & Berg, J. (2018). From wind to loads: wind turbine site-specific load estimation with surrogate models trained on high-fidelity load databases. Wind Energy Science, 3(2), 767–790. https://doi.org/10.5194/wes-3-767-2018

[4] Dimitrov, N. K. (2019). Surrogate models for parameterized representation of wake-induced loads in wind farms. Wind Energy, 22(10), 1371–1389. https://doi.org/10.1002/we.2362

[5] Yan, C., Pan, Y., & Archer, C. L. (2019). A general method to estimate wind farm power using artificial neural networks. Wind Energy, 22(11), 1421–1432. https://doi.org/10.1002/we.2379

[6] Larsen, G. C., Madsen Aagaard, H., Bingöl, F., Mann, J., Ott, S., Sørensen, J. N., … Mikkelsen, R. (2007). Dynamic wake meandering modeling. Risø National Laboratory.

[7] Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. Reliability Engineering and System Safety, 93(7), 964–979. https://doi.org/10.1016/j.ress.2007.04.002

[8] Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Mathematics and Computers in Simulation, 55(1-3), 271–280. https://doi.org/10.1016/S0378-4754(00)00270-6

[9] IEC (2005) International Standard IEC61400-1: wind turbines—part 1: design guidelines, Third edition.

[10] B.J. Jonkman and M.L. Buhl, Jr. (2006) TurbSim User's Guide. NREL Technical Report TP-500-39797

[11] Mann, J. (1994). The spatial structure of neutral atmospheric surface-layer turbulence. Journal of Fluid Mechanics, 273, 141–168. https://doi.org/10.1017/S0022112094001886

[12] Larsen, T. J., & Hansen, A. M. (2007). How 2 HAWC2, the user's manual. Risø National Laboratory.

[13] Berg, J., Natarajan, A., Mann, J., & Patton, E. G. (2016). Gaussian vs non-Gaussian turbulence: impact on wind turbine loads. Wind Energy, 19(11), 1975–1989. https://doi.org/10.1002/we.1963

[14] Hiperwind Deliverable 2.3: Environmental joint probability distributions and uncertainties. https://www.hiperwind.eu/publications (accessed March 2022).

[15] Yassin, K., Helms, A., Moreno, D., Kassem, H., Höning, L., and Lukassen, L. J.: Applying a Random Time Mapping to Mann modelled turbulence for the generation of intermittent wind fields, Wind Energ. Sci. Discuss. [preprint], https://doi.org/10.5194/wes-2021-139, in review, 2021

[16] Hannesdóttir, Á., Kelly, M. C., & Dimitrov, N. K. (2019). Extreme wind fluctuations: Joint statistics, extreme turbulence, and impact on wind turbine loads. Wind Energy Science, 4(2), 325–342. https://doi.org/10.5194/wes-4-325-2019

[17] Kelly, M., Andersen, S. J., & Hannesdóttir, Á. (2021). Statistical impact of wind-speed ramp events on turbines, via observations and coupled fluid-dynamic and aeroelastic simulations. *Wind Energy Science*, *6*(5), 1227-1245. https://doi.org/10.5194/wes-6-1227-2021.

[18] Dimitrov, N. K., & Natarajan, A. (2017). Application of simulated lidar scanning patterns to constrained Gaussian turbulence fields for load validation. Wind Energy, 20(1), 79–95. https://doi.org/10.1002/we.1992

[19] https://pypi.org/project/hipersim/ , accessed March 2022.

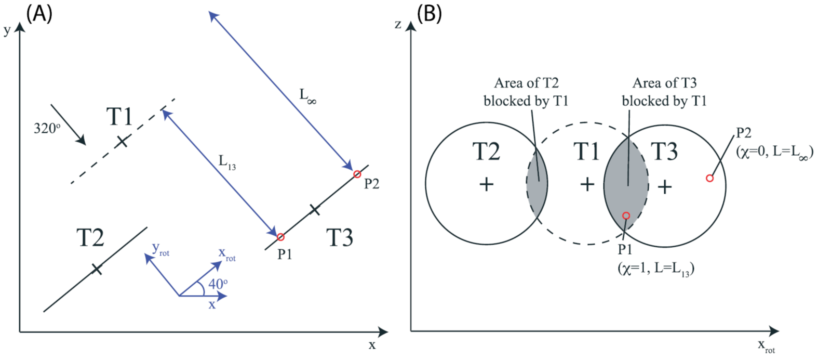
3. Wind farm wake effect parameterization using layout encoding

3.1. Proposed layout parameterization approaches

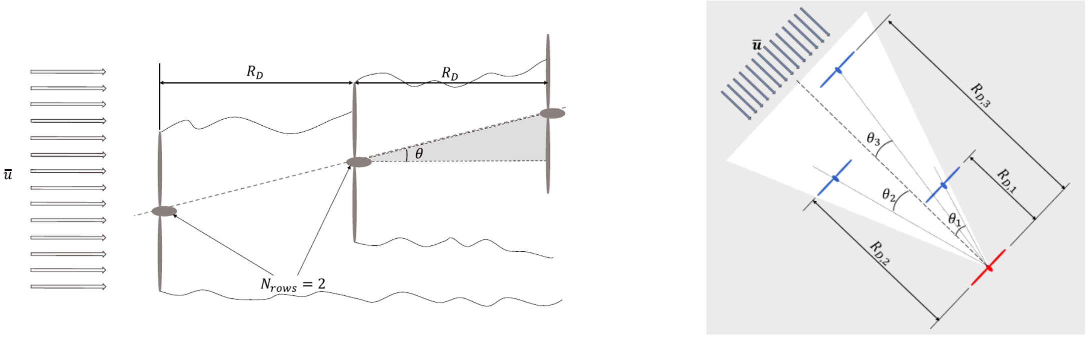
In this part of the report, we examine the following parameterization approaches based on manually extracted layout features:

- 3-parameter approach proposed by Dimitrov (2019) [1]
- 2-parameter approach proposed by Yan et al (2019) [2]
- 2-parameter-per-turbine approach proposed by Dimitrov in (2021) [3]

The 3-parameter approach is based on the concept of considering groups of aligned upwind turbines as one row of turbines. The 3 parameters include row spacing, number of upwind turbines in the row, and the relative wake angle for the row (the angle between the row alignment and the wind direction). This approach is suitable for regular wind farm layouts consisting of straight rows or a regular grid of turbines. For a given turbine and ambient wind direction, the upwind turbines can form a few rows. In each row, the turbines are considered to have equal spacing. The chosen row of upwind turbines to represent the wake-induced effects have the minimum distance to the disturbed turbine. While this approach has shown great success in surrogate modelling of turbine power and loads in wind farms with regular layouts [1], its definition is not suitable for irregular wind farm layouts where rows cannot be clearly identified.



(a) Blockage-based features (figure reproduced from [2])



(b) Row-based features
(figure reproduced from [1])

(c) Turbine-wise features
(figure reproduced from [3])

Figure 4 Three approaches to define features to describe wake sources

The 2-parameter approach of Yan et al. [2] is based on the concept of blockage. The 2 parameters include the blockage ratio and the blockage upwind distance (Figure 4, top center plot). Yan et al. (2019) have used these two parameters to build surrogate models for estimating the power generation from the entire wind farm. However, with the present study we found that Yan's two-parameter model is less capable of predicting the power or loads of individual turbines, compared to its ability to predict an entire farm's power. This may be explained by the fact that the concept of blockage ignores wake expansion, and thus some upwind turbines whose wake can in fact disturb a turbine. Improvements have been made in relevant studies [4] to include more geometric quantities, such as the crosswind blockage distance (an aggregate measure of the relative cross-wind position of upwind turbines), and to consider the expansion of wakes. This 2-parameter approach and its modifications are by definition not limited to regular wind farm layouts; they can also be applied to irregular wind farm layouts.

The 2-parameter "per-turbine" approach (Dimitrov, 2021 [3]) uses 2 geometric parameters (e.g. one relative wake angle and one turbine-to-turbine distance) to describe each chosen upwind turbine. Turbines are considered as potential wake sources only if they are located within a sector of ± 16deg with respect to the incoming wind direction. In practice, a limit is also needed on the number of chosen upwind turbines in order to have a fixed number of variables as model inputs. In this study, up to 20 of the closest upwind turbines are chosen, according to turbine-to-turbine distance. If for an actual turbine and wind direction under consideration there are fewer than 20 turbines located in the upwind sector, the redundant geometric parameters are by default set to zeros so that we always have 40 (20 distances and 20 relative angles) model input variables. This approach by definition generally applies to any wind farm layout, including both regular and irregular.

Besides manually defined features, this project also studied a set of deep learning approaches, such as autoencoder, for the purpose of feature learning and dimension reduction. The autoencoder is a special example of a Feedforward Neural Network (FNN) where one of the hidden layers has significantly smaller number of units than the number of input variables. In a typical application, the model is trained to reproduce its inputs with the best possible accuracy. This information needs to pass through the "bottleneck" layer and as a result, the model is forced to condense the most important information in few quantities represented by latent variables (the activation outputs of the bottleneck layer). In the present study, the autoencoder architecture is used in pre-processing models that reduce the dimensionality of the farm-related inputs down to a few latent variables. These latent variables are used further as inputs to another surrogate model for loads or power prediction. Three autoencoder implementations are used:

- a sparse autoencoder for encoding all the turbine positions (including both upwind and downwind turbines) in Cartesian coordinate system relative to the considered turbine and the ambient wind direction.
- an autoencoder for encoding the positions of up to N closest upwind turbines within the relative angle limit (e.g. 16 degrees)
- a deep-learning neural network approach, which has encoder and decoder that are similar to the above approach. However, the target output during training is the power accounting for the wake-induced effects.

The first autoencoder approach encodes all the turbines positions relative to a considered turbine and a given wind direction. The resulting latent variables show a smooth variation against the wind direction for a considered turbine. The reconstructed wind farm layout is similar to the input wind farm layout but with a different row spacing. This autoencoder approach is limited to one wind farm layout, which means the learned features depend on the difference of the turbine positions in a wind farm. However, its architecture is extremely simple and thus fast to train.

The second autoencoder encodes the positions of up to N closest upwind turbines within a relative angle limit (e.g. 16 degrees). Despite testing a variety of architectures, this autoencoder does not provide desirable performance in reconstruction of the input turbine positions and the later modelling of the power and loads. The wind farm layouts for training the autoencoder is randomly generated, including 50 regular wind farm layouts and 50 irregular wind farm layouts.

The third deep learning approach retains the general model architecture from the second approach, but adds one output layer and targets the prediction of the power output of a dummy turbine rather than recovering the inputs. Due to this feature, the model is no longer an autoencoder, because it maps the upwind turbine positions to the turbine's power accounting for wake effects. By including wake-induced effects in the output, this deep learning approach to some extent ensures that the learned features represented by the latent variables are related to the wake-induced effects. For this deep learning approach, 10 regular wind farm layouts and 10 irregular wind farm layouts are used in the training. Then the trained encoder is applied to a new wind farm site to test its applicability. The turbine in the training is also a mock turbine different from the turbine in the new wind farm site. The learned features in this approach are not limited to one wind farm layout, and can be applied to general wind farm layouts.

Note that for all these three latent-variable neural network models, only the trained encoder is directly used in the modelling of the turbine's power or loads without re-training. Table 1 summarizes the architecture of the six neural network models.

Table 1 Architecture of the six neural network models

Model	Architecture	
	Layer sizes	Activation function
1st feature-based NN	4-20-20-20-1	Relu-Relu-Relu-Linear
2nd feature-based NN	3-20-20-20-1	Relu-Relu-Relu-Linear
3rd feature-based NN	41-40-40-40-1	Relu-Relu-Relu-Linear
1st latent-variable NN	11-20-20-20-1	Relu-Relu-Relu-Linear
2nd latent-variable NN	11-20-20-20-1	Relu-Relu-Relu-Linear
3rd latent-variable NN	5-40-40-40-1	Relu-Relu-Relu-Linear

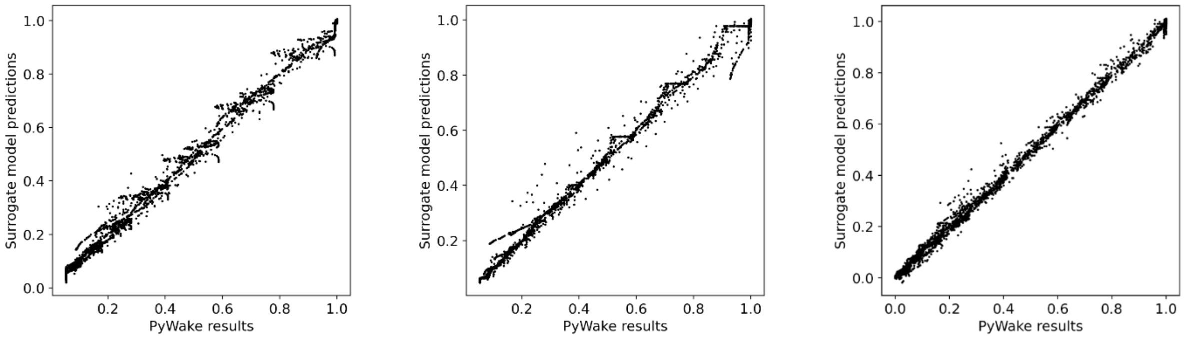
In the above, Relu stands for "Rectified linear unit", and linear stands for an activation function with the formula $a(x) = wx + b$, where w and b are real-valued weights. The latter function is appropriate as an output layer for a Neural Network acting as a regression model.

3.2. Assessment of layout parameterization approaches using Pywake

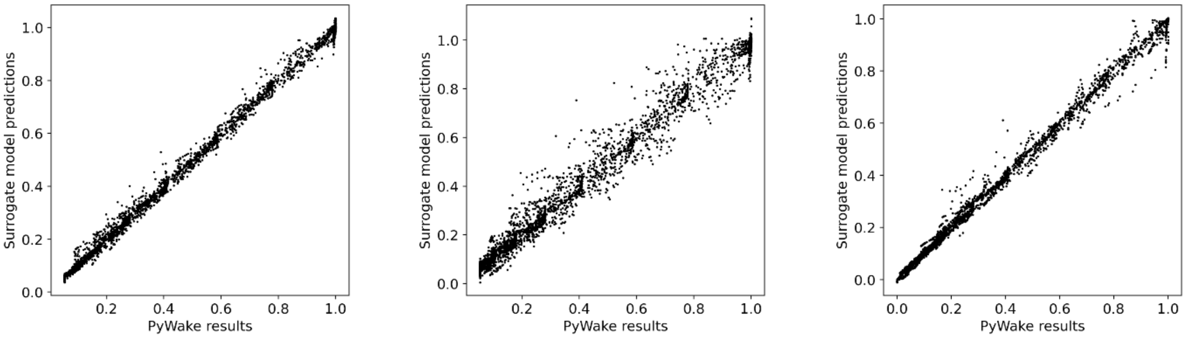
These approaches were applied to an operational offshore wind farm. The aim is to compare the effectiveness of these manually or automatically extracted features to represent the wake-induced effects on power or loads. It is reasonable to think that if the features can be successfully employed to parameterize wake-induced effects on power, they can also be used to parameterize the wake-induced effects on loads. Below, we assess these features by using data on turbine's power generated from PyWake.

PyWake [5] is a wind farm simulation toolbox developed at DTU Wind Energy. It computes each turbine's power production in a wind farm based on engineering wake-deficit models and turbulence models. Here we use the wake deficit model by Zong and Porté-Agel (2020) [6], the turbulence model by Crespo and Hernández (1996) [7], and wake superposition is obtained using the squared sum of individual deficits [8].

Figure 5 compares the correlation of the simulated and estimated turbine power for the six neural network models. It can be seen that three models correlate well with the simulated data, including the 3rd feature-based neural network model, the 1st and 3rd latent-variable neural network models.



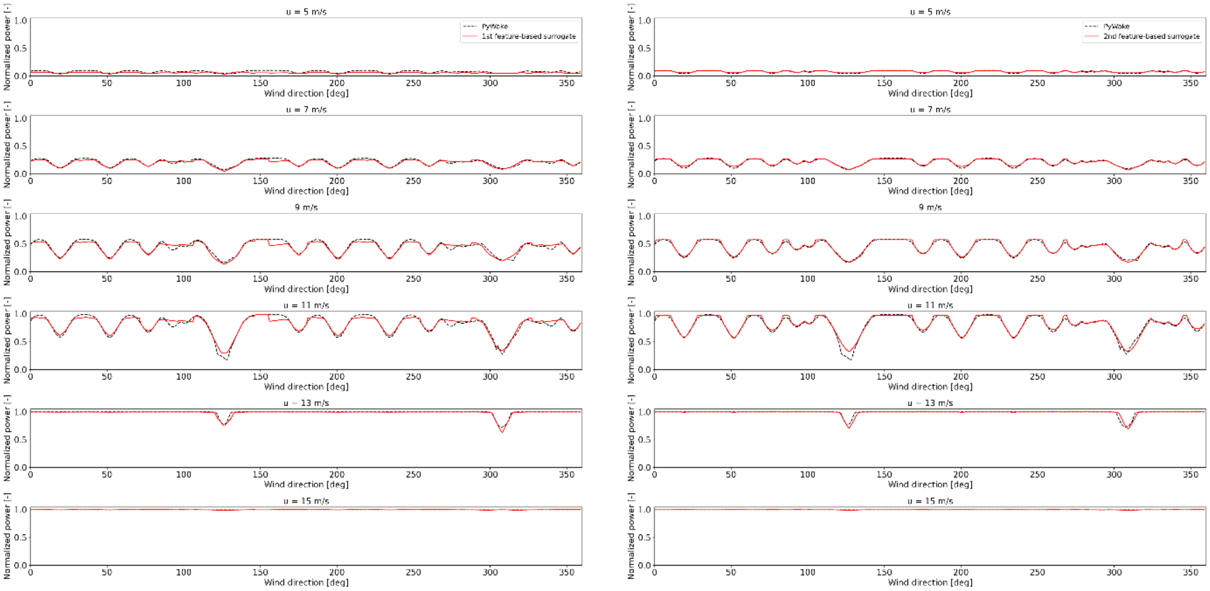
(a) 1st feature-based NN (b) 2nd feature-based NN (c) 3rd feature-based NN



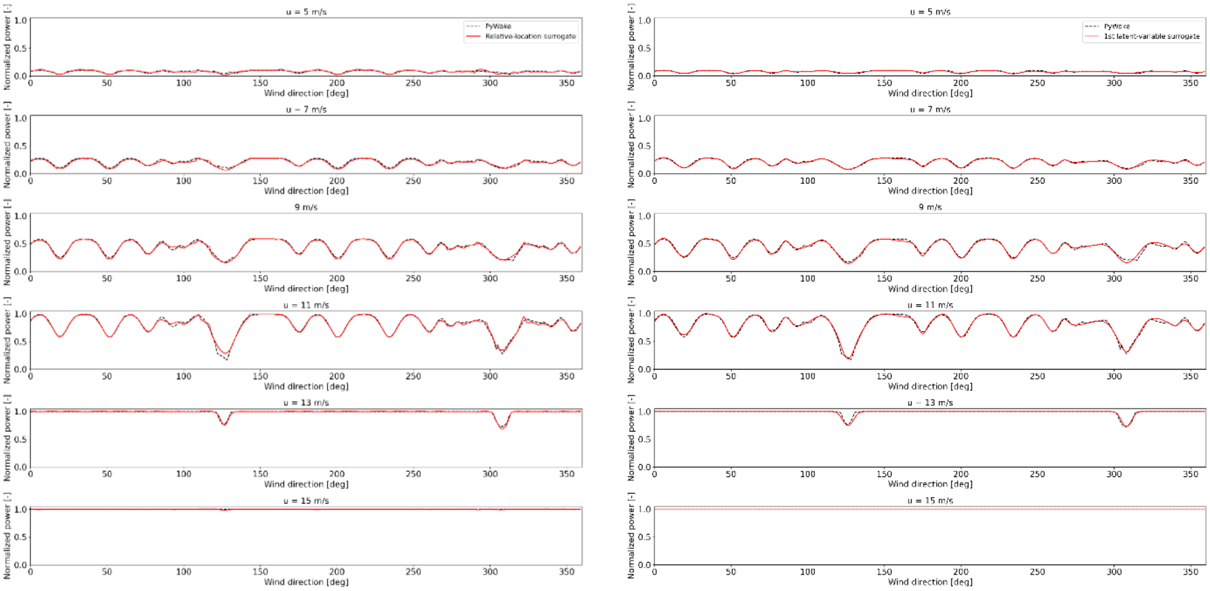
(d) 1st latent-variable NN (e) 2nd latent-variable NN (f) 3rd latent-variable NN

Figure 5 Comparison of the six neural network models in terms of correlation between simulated and estimated data. The plotted data is for one turbine in the center of the offshore wind farm with different ambient wind speeds and directions

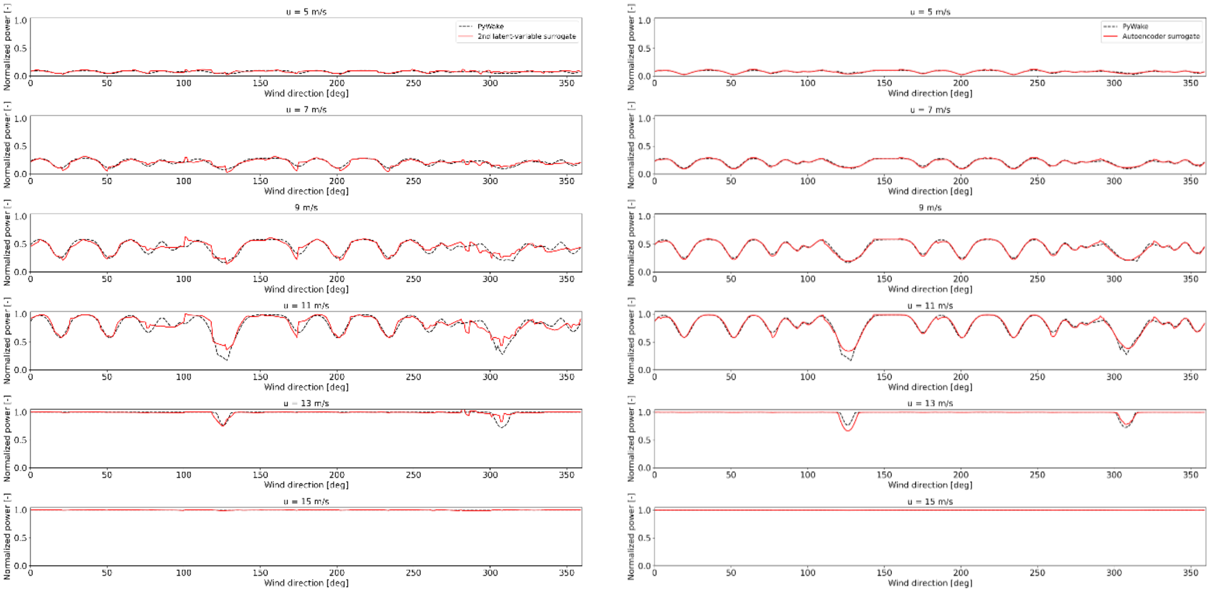
The following Figure 6 compares the six neural network models in terms of the variation of the turbine power with the ambient wind direction. This kind of plot can clearly show the wake-induced effects, and therefore allows us to gain a physical interpretation of the model's performance. It can be seen that all the six neural network models are able to capture the locations of the wake sources. However, they may over- or under-estimate the reduction of turbine power.



(a) 1st feature-based surrogate (b) 2nd feature-based surrogate



(c) 3rd featured-based surrogate (d) 1st latent-variable surrogate



(e) 2nd latent-variable surrogate (f) 3rd latent-variable surrogate

Figure 6 Comparison of the six neural network models for one turbine and different ambient wind speeds: 5 m/s , 7 m/s, 9 m/s, 11 m/s , 13 m/s, 15 m/s.

The following table shows a comparison of the performance of the six neural network models in terms of correlation and errors. The 1st latent variable neural network model has both the highest correlation and the lowest error between simulated and estimated data. This implies that the 1st latent variable NN is suitable for parameterization of the wake-induced effects when the application is related to only one wind farm. Both the 3rd feature-based neural network and the 3rd latent-variable-based neural network show high and comparable performance in terms of correlation and errors. These two models can be applied to general wind farm layout. The 3rd latent-variable-based neural network model requires a smaller number of input variables.

Table: Comparison of the performance of modelling wake-induced effects on power by using different features

Model	**R-squared**	**Mean bias**	**MAE**	**RMSE**
1st feature-based surrogate	0.9983	0.2142%	0.7071%	1.5449%
2nd feature-based surrogate	0.9986	0.0193%	0.5042%	1.3843%
3rd feature-based surrogate	0.9988	-0.0350%	0.4626%	1.2219%
1st latent-variable surrogate	0.9993	0.0078%	0.3234%	0.9976%
2nd latent-variable surrogate	0.9955	-0.2377%	0.9961%	2.4946%
3rd latent-variable surrogate	0.9988	0.1270%	0.4408%	1.2212%

3.3. Load surrogates with the DWM model

Up until now, we have assessed multiple layout parameterization approaches based on their capability of predicting wind farm power output. When expanding the surrogate-based approach to aeroelastic loads, several additional factors need to be considered:

1) The computational expense of running aeroelastic simulations is orders of magnitude higher than with the quasi-static wake models such as PyWake;
2) There is realization-to-realization uncertainty due to the random turbulence fields used in the simulations; hence multiple realizations at each sample point are necessary in order to reduce this uncertainty.
3) The effect of ambient turbulence and wind shear is potentially more pronounced than with the quasi-static wake models, both due to modelling differences (e.g., the wake meandering driven by turbulence) and due to higher sensitivity of the loads to the variation in the wind field — the power is mostly affected by the total energy over the rotor area, while the loads are subject to cyclic variations when the individual blades sample the non-uniform wind field.

Study setup

Taking into account the above factors and limitations means that a new design of experiment (DoE) needs to be developed. We need at least three ambient variables (wind speed, turbulence intensity, wind shear), and together with the wake-related variables the number of input dimensions is too high to make a grid-like DoE feasible. Therefore, we use a space-filling design based on a Halton quasi-random sequence [9] with a total of 30000 sampling points. The DoE includes five variables with the distributions defined in Table 2. The distribution bounds are adapted from [10]. The wake conditions are not explicitly included in the design, however based on the wind direction and the turbine number, wake-related variables can be determined for any of the parameterization approaches described in Section 3.1. The sample generation is done in a two steps process: 1) a set of uniform, i.i.d. variables ranging from 0 to 1 is drawn from the Halton quasi-random sequence, and 2) the physical values with the prescribed conditional distributions are obtained by a Rosenblatt transformation. Note that the Halton sequence is only used for the first four variables in Table 2, and the wind turbine index numbers are drawn independently from the other variables by directly generating random integers from 1 to 27.

Table 2 Variable distribution definitions for aeroelastic load simulation inputs

Variable	Notation	Distribution
Wind direction	θ	Uniformly distributed in the range $(0,360)°$, discretized in steps of $2°$
Ambient wind speed	u	Truncated Weibull distribution in the range $[4,25]$m/s, with parameters $A = 15$ and $k = 1$. This is not equivalent to the site-specific conditions but the modified distribution is chosen in order to fill the space better.
Turbulence (wind speed standard deviation	σ_u	Uniformly distributed, with bounds conditional on u: - $\sigma_{u,LB} = 0.025u$ - $\sigma_{u,UB} = 0.16(0.012u^2 + 0.75u + 5.6)$
Wind shear exponent	α	Uniformly distributed, with bounds conditional on u: - $\alpha_{LB} = 0.05 - 0.05\left[1 - \left(0.4\log\left(\frac{0.5D}{z_{hub}}\right)^2\right)\right]\left(\frac{u_{cut-out}}{u}\right)$

		- $\alpha_{UB} = 0.15 + 0.1\left(\frac{D}{z_{hub}}\right)\left(\frac{u_{cut-out}}{u}\right)$
Wind turbine number	T	Uniform random integers from 1 to 27

In the above, LB and UB stand for lower bound and upper bound respectively, z_{hub} is the wind turbine hub height, D is the rotor diameter, and $u_{cut-out}$ is the cut-out wind speed of the wind turbine.

The manifold of input conditions focuses on just a few environmental variables that are deemed most important for the loads. Air density is not considered because the effect of changing air density is similar in nature to changing the wind speed (the effect of changing air density on the aerodynamic forces is equivalent to that of changing the square of the wind speed). Veer is not considered due to the lack of measurements near or above hub-heights at the Teesside wind farm, needed to show the veer in the upper rotor half; it is not possible to validate the effect of veer with measurements. Turbulence spectral parameters are not considered due to the lack of high-frequency data to estimate them. Instead, the standard values of the Mann model parameters are used (length scale L=29.4 m, anisotropy parameter Γ=3.9). For a given value of L and Γ, the third Mann model parameter, $\alpha\varepsilon^{2/3}$, is proportional to the wind field variance, hence it has

The sample space as defined above is used as input to aeroelastic load simulations using a 2.3MW wind turbine model implemented in the HAWC2 tool [11]. In order to account for the realization-to-realization uncertainty, six aeroelastic load simulations are carried out at each sampling point – a total of 180,000 simulations. Each simulation has 10 minutes duration. The following output channels from the load simulations are collected:

1) damage-equivalent fatigue loads (DELs) for 6 signals: tower base fore-aft ($M_{x,TB}$) and side-to-side ($M_{y,TB}$) bending moments, blade root flapwise ($M_{x,B}$) and side-to-side ($M_{y,B}$) bending moments, tower top yaw (torsion) moment $M_{z,TT}$, and shaft torsion moment, $M_{z,S}$;
2) mean values for the electrical power, blade pitch, and rotor rotational velocity.

The model training and evaluation dataset is obtained by averaging the results from all 6 realizations at each sampling point. As a result, we get 30000 input-output pairs. Based on the outcome of the comparisons shown in Section 3.2, the wake parameterization approach based on an autoencoder with layout-derived features and with power output as target (referred to as 3rd latent-variable NN) is chosen. This approach shows good performance while it can also be applied on an arbitrary wind farm without retraining. Implementing the autoencoder model as means of wake parameterization leads to the workflow defined in Figure 7. Then, computing the surrogate model inputs at a given sample point is with the following procedure: 1) based on the wind direction, the number of wind turbine of interest, and the wind turbine coordinates for the full wind farm, compute the layout-derived features that serve as input to the autoencoder. These are the upwind and cross-wind distances for the 20 closest turbines upwind within a sector with $\pm$16deg width. 2) run the autoencoder to obtain the latent variables encoding the wind farm layout features. In the example shown below, we use an autoencoder with 4 output features. 3) Assemble a single array from the latent variables and the ambient variables. These are the inputs to the actual load & power surrogate model. In the present case, the total number of inputs is 7: 3 ambient variables and 4 latent variables describing the wake interactions. The entire procedure of computing the surrogate model inputs has been automated and is included as a script in the Hipersim software tool.

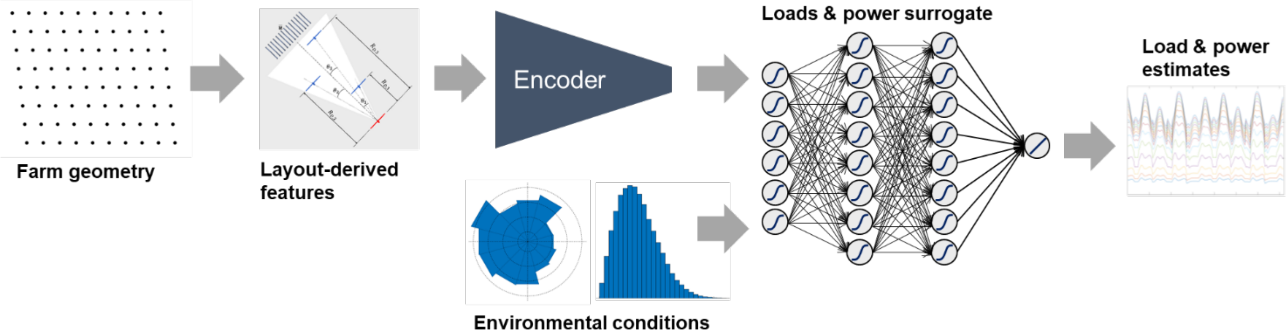


Figure 7 Sketch of load surrogate modelling workflow

We use the Hipersim neural network toolbox to train a surrogate model using 90% of the available data (the first 27000 points of the Halton sequence). The remaining 3000 data points are retained as a test set for model performance evaluation. The training algorithm is adaptive moment estimation (Adam [12]) with minibatch size of 1000, and learning rate of 0.001. Other hyperparameters, such as model size and activation function, were chosen based on multiple trial runs with different setups. The final setup chosen is two hidden layers with 60 units each, and a ReLU activation function.

Results

As mentioned earlier, we focus the analysis of the load predictions on turbine #14 in the Teesside wind farm as it has complex inflow conditions from multiple wind directions. To get first insight at these wake effects, we plot the values of the wake-encoding latent variables for a 0–360° sweep of wind directions for wind turbine #14. This is shown in Figure 8. The peaks in some latent variables clearly indicate the directions where wake disturbances are present. However, it is not possible to directly ascribe physical meaning to any of the variables, and it is also not all variables that behave in a clearly explainable way. However, the directions where the turbine is presumably in the free wind (around 40°, 220° and 330°) are clearly recognizable as all latent variables have step changes in these regions.

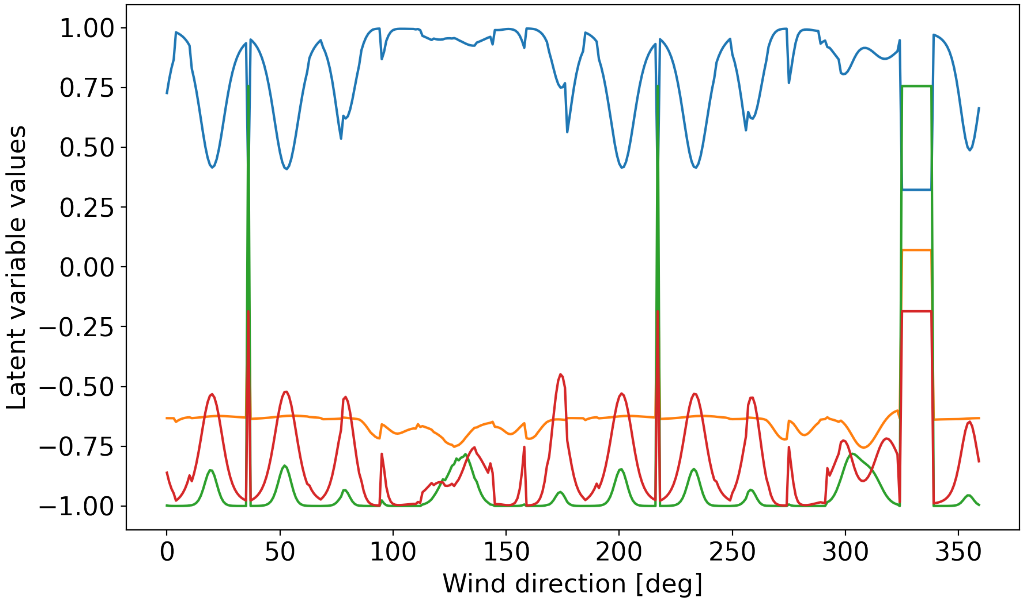


Figure 8 Wake-encoding latent variables as function of wind direction, T14 at Teesside

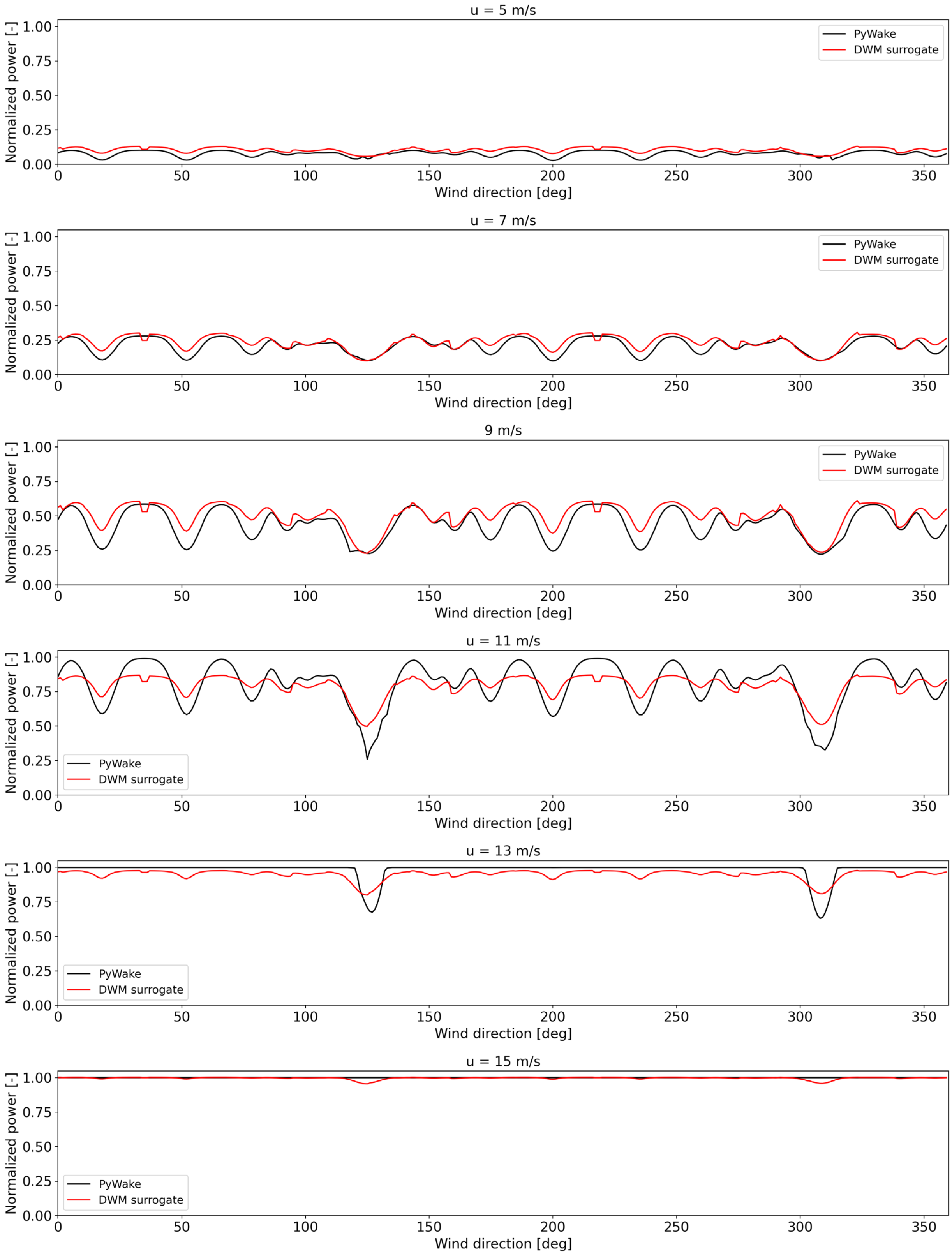


Figure 9 Mean power output predictions – PyWake vs. surrogate model trained on Hawc2 simulations.

The currently used surrogate approach showed very good performance in predicting power outputs as computed with PyWake. It is relevant to test the performance of the surrogate trained on Hawc2 aeroelastic simulations against the same reference PyWake case. Figure 9 shows the power output predictions of the aeroelastic simulation surrogate against PyWake predictions, for turbine #14 at Teesside. For most wind speeds the Hawc2-based surrogate shows similar good performance as the PyWake-based surrogates discussed in Section 3.2. At medium wind speeds (11-13m/s) there are bigger inconsistencies. The main patterns of wake deficits are captured; however the depths of the deficit do not agree. It is also visible that the power at wake-free or almost wake-free conditions is not the same between the Hawc2 surrogate and PyWake. This leads to the conclusion that the main contributor for the discrepancy may not be the surrogate model inaccuracy, but rather differences in the power curves simulated by PyWake and by Hawc2 (which can be caused by the presence of wind turbine controller in Hawc2, as well as different sensitivity to the effect of turbulence for operation close to the saturation point on the power curve). In order to test this assumption as well as make a more in-depth assessment of the model performance, we need a comparison between the surrogate model predictions and reference aeroelastic simulations, which is described below.

The test data set of 3000 points sampled with a space-filling technique can be used to evaluate the general performance of the surrogate models over the full range of conditions. We evaluate this performance using two standard measures – normalized root mean squared error (NRMSE), and coefficient of determination, R-squared. The test metrics for all considered load quantities are listed in Table 3.

Table 3 Model performance metrics evaluated on a test data set of 3000 points

Signal name	NRMSE	R-squared
Mudline Mx DEL	0.163	0.892
Mudline My DEL	0.165	0.933
Tower base Mx DEL	0.204	0.834
Tower base My DEL	0.167	0.935
Tower top Mx DEL	0.122	0.961
Tower top My DEL	0.172	0.917
Tower top Mz DEL	0.122	0.962
Blade root Mx DEL	0.123	0.938
Blade root My DEL	0.022	0.917
Shaft Mz DEL	0.213	0.9
Rotor RPM	0.046	0.942
Blade pitch	0.113	0.993
Mean power	0.093	0.973
Mudline Mx Max	0.069	0.952
Mudline My Max	0.125	0.947
Tower base Mx Max	0.073	0.951
Tower base My Max	0.12	0.947
Tower top Mx Max	0.511	0.912
Tower top My Max	0.052	0.985
Tower top Mz Max	0.178	0.926
Blade root Mx Min	0.05	0.969
Blade root My Min	0.043	0.929
Shaft Mz Mean	0.171	0.936

Due to the unstructured (not aligned) nature of the space-filling design used, it is not possible to make a direct visual comparison for some representative cases such as those shown in Figure 9. Instead, we carry out an additional set of simulations for selected cases with turbine #14, i.e., varying the wind direction from 0 to 360 deg, while keeping all other conditions constant. In order to evaluate the realization-to-realization uncertainty, 12 different realizations are simulated at each input condition. The mean and the standard deviation from the 12 realizations are then computed. The following figures show a comparison of the outcomes from these cases to the surrogate model predictions. In Figure 10 that compares the power output, the PyWake results are also included. Figure 11 to Figure 15 show a comparison of damage equivalent fatigue loads (DEL) prediction – for mudline (sea bottom) fore-aft bending, tower top fore-aft bending, tower top side-side bending, tower top torsion, and blade root flapwise bending respectively. Finally, Figure 16 and Figure 17 show the prediction of extremes (the maximum attained value over a 10-minute period), for the fore-aft bending moment and the blade root flapwise bending moment, respectively.

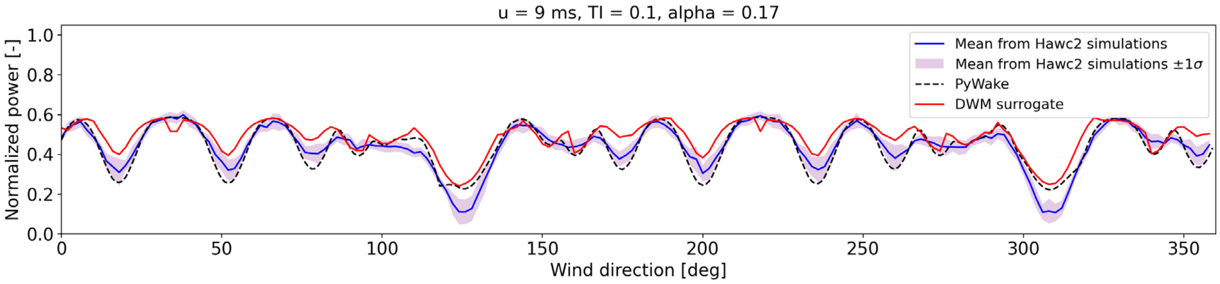


Figure 10 Comparison of ***mean power output*** *results from Hawc2 simulations vs. predictions from a surrogate model*

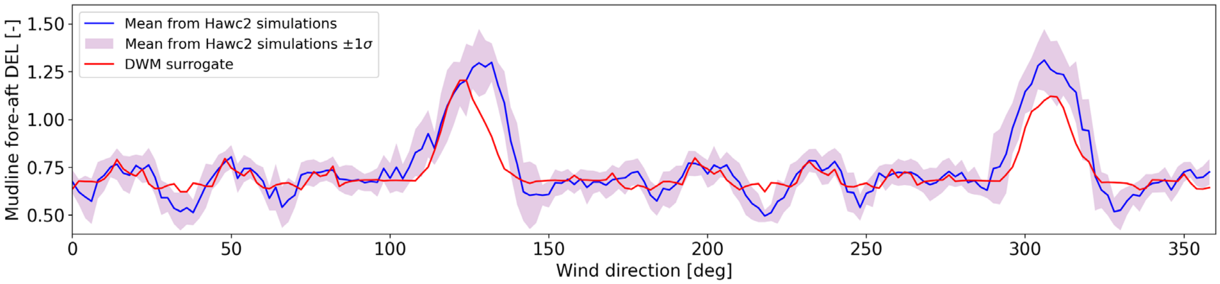


Figure 11 Comparison of the ***mudline fore-aft bending DEL*** *results from Hawc2 simulations vs. predictions from a surrogate model*

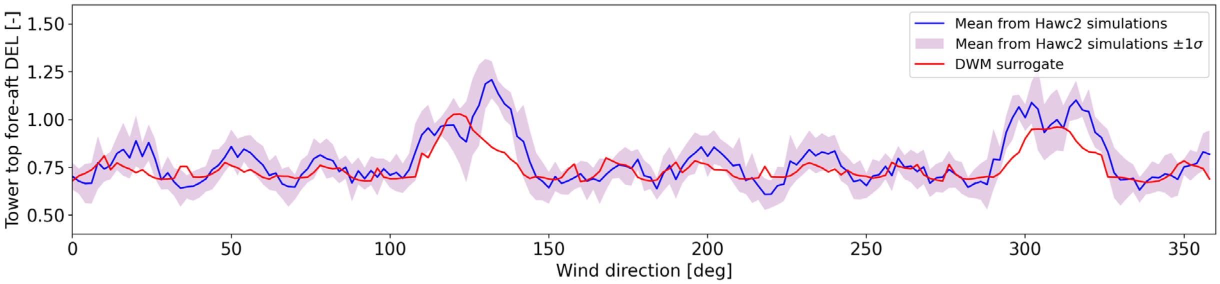


Figure 12 Comparison of ***tower top fore-aft bending DEL*** *results from Hawc2 simulations vs. predictions from a surrogate model*

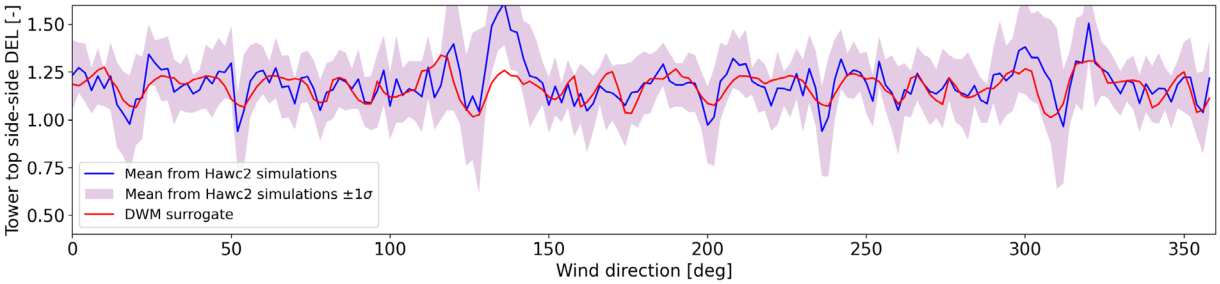


*Figure 13 Comparison of **tower top side-side bending DEL** results from Hawc2 simulations vs. predictions from a surrogate model*

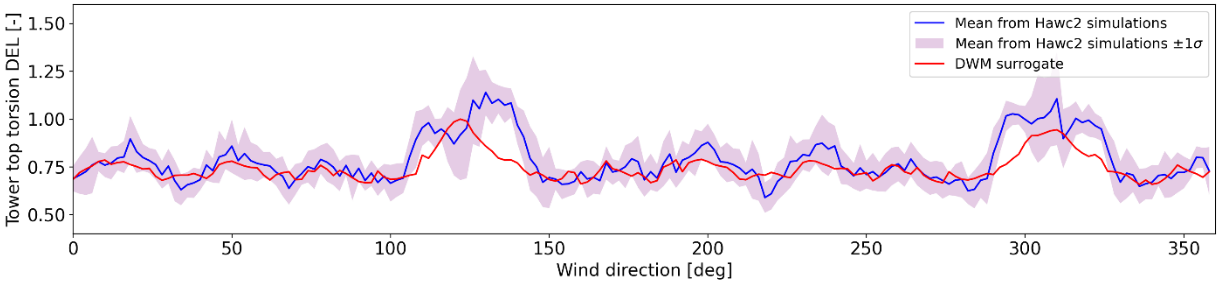


*Figure 14 Comparison of **tower top torsion DEL** results from Hawc2 simulations vs. predictions from a surrogate model*

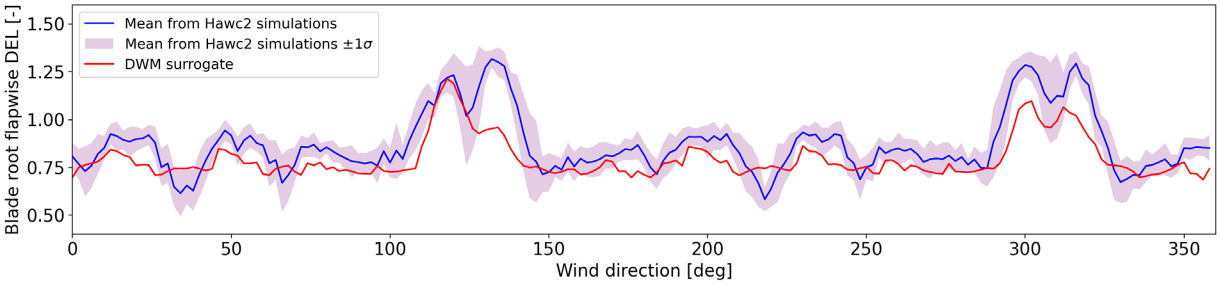


*Figure 15 Comparison of **blade root flapwise bending DEL** results from Hawc2 simulations vs. predictions from a surrogate model*

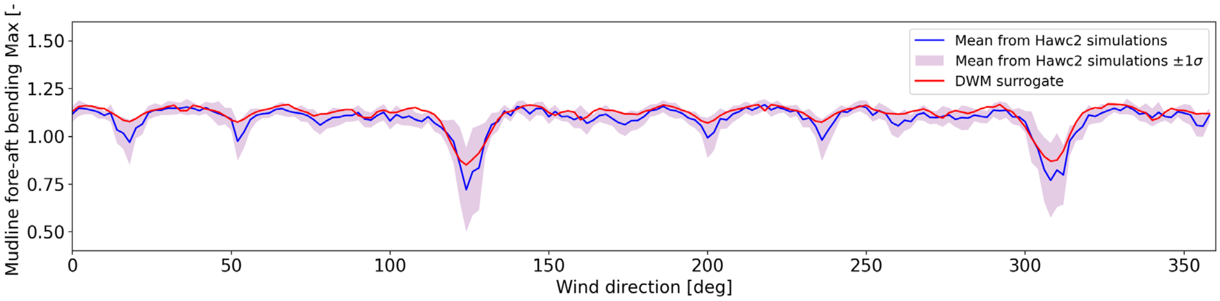


*Figure 16 Comparison of **mudline fore-aft bending MAX** results from Hawc2 simulations vs. predictions from a surrogate model*

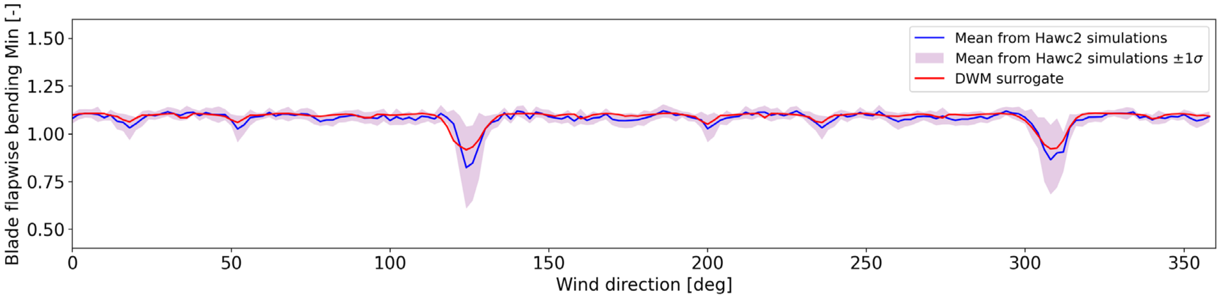


Figure 17 Comparison of ***blade root flapwise bending MAX*** *results from Hawc2 simulations vs. predictions from a surrogate model*

The tendencies seen in the above observations agree well with those apparent in the performance metrics listed in Table 3. We summarize the conclusions from the load surrogate modelling in wake conditions as follows:

- Overall, the currently investigated load surrogates capture the main trends in the load behavior and the impact of wakes, however the performance is somewhat worse than for the case of predicting static wake model outputs such as with PyWake. Since the main critical difference between the PyWake data and the load simulation data is the realization-to-realization uncertainty, it is likely that this is one of the main reasons for the model performance loss. This may be rectified with increasing the number of sample points, or increasing the number of realizations per sample – with both approaches expected to have a similar effect.
- Fatigue load DELs and mean values such as mean power output show significantly stronger wake signatures, compared to extremes. For the extremes we even see that the wakes lead to load reduction in the case at 9m/s – likely due to the reduction in the equivalent wind speed in the wake. The high sensitivity of DEL can be at least partly explained by the exponentiation of load cycles with the Wöhler exponent, which amplifies any variations in the signals.
- The performance of the surrogates is affected by the significance of wake effects: for the channels with less wake impact, such as the extremes, the performance is consistently higher.

Global sensitivity analysis

We carry out variance-based sensitivity analysis in order to study the relative importance of the ambient conditions compared to the wake effects for loads and power in wind farms. The main indicator used is the total-effect Sobol indices that indicate the total contribution of a given input variable to the variance of a model output [13]. Utilizing the computational efficiency of the load surrogate models, the Sobol indices are computed using a Monte Carlo simulation following the method proposed by [14]. The approach is briefly described as follows:

1) Generate two input samples of same size, **A** and **B**, each with dimensions $N \times m$, where N is the number of Monte Carlo samples and m is the input dimension of the model;
2) For $j = 1 \dots m$, create input matrices $\mathbf{AB}_j$ such that the j^{th} column in $\mathbf{AB}_j$ is equal to the j^{th} column in **B**, while the remaining columns are equal to those in **A**;
3) Evaluate the model function $g()$ for both input matrices and over all samples: $g\left(\mathbf{A}^{(i)}\right)$ and $g\left(\mathbf{AB}_j^{(i)}\right)$ for $i = 1 \dots N$ and $j = 1 \dots m$
4) Compute the total-effect indices with the following estimator function:

$$S_{T,j} = \frac{\frac{1}{2N}\sum_{i=1}^{N}\left(g\left(\mathbf{A}^{(i)}\right) - g\left(\mathbf{AB}_j^{(i)}\right)\right)^2}{\mathrm{Var}(\mathbf{y})}$$

where $\mathrm{Var}(\mathbf{y})$ is the vector of variances of the output variables.

We apply this method with a Monte Carlo simulation based on turbine #14 in the Teesside wind farm, with the first 4 variables from Table 2 $(\theta, u, \sigma_u, \alpha)$ as random inputs. As the model variance and subsequently the Sobol indices may depend on the distribution of input conditions, we need to choose a suitable reference distribution for the input variables. The sampling distribution given in Table 2 is aimed at best surrogate model training and is not representative of real-world conditions. We therefore choose a reference distribution that is consistent with the IEC 61400-1 standard definition of wind conditions: the wind speed is Weibull distributed with parameters $A = 11.28$ m/s and $k = 2$ (wind class I), turbulence is log-normally distributed and conditional on wind speed with the formula $E(\sigma_u|u) = 0.12(0.75u + 3.8\ \mathrm{m/s})$; $Var(\sigma_u) = 0.12 \cdot 1.4\ \mathrm{m/s}$ (corresponding to turbulence class C), and the wind shear distribution is based on [15] with $E(\alpha|u) = 0.088(\log u - 1)$ and $Var(\alpha|u) = 1/u^2$ following [16]. The wind direction distribution is considered uniform across the range $(0{,}360)$ degrees.

Note that although the surrogate model uses wake parameterization variables as inputs, it is not straightforward to directly compute Sobol indices for these variables since there is no explicit formulation for their (joint) probability distribution. Instead, since we know that all the wake parameterization variables are dependent on wind direction, we will consider the wind direction as a variable summarizing the effect of wake interactions.

A Monte Carlo (MC) simulation with the above procedure is carried out for surrogate models covering all of the previously considered output channels. As a first check, we plot the convergence history of the total-effect Sobol indices over the 10000 Monte Carlo samples, which is shown in Figure 18 for the power output. The chosen MC sample size leads to good convergence of the Sobol index estimates.

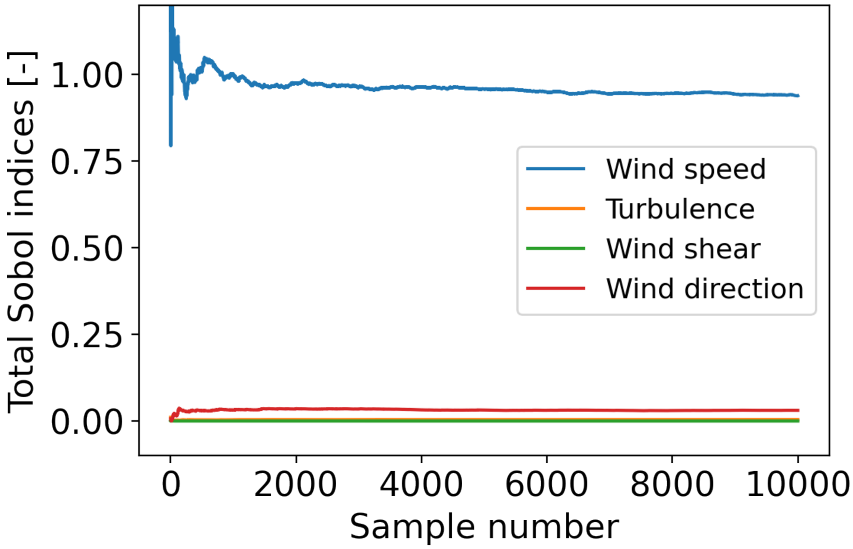


Figure 18 Total effect Sobol indices for power output

Next, the procedure is repeated for all channels considered. The results for S_T for all signals are listed in Table 4. Note that the sum of the total effect Sobol indices can exceed 1 as they also account for interactions between input variables (meaning that the effect of an interaction term is ascribed to all variables included in the term). Theoretically each individual index should be less than 1, however as the indices in Table 4 are derived from a random simulation they are approximate and uncertain – and in a few cases exceed 1. For most channels, the wind speed is the dominant factor, however for a few, mainly fatigue channels, the turbulence has also a significant effect.

Table 4 Total-effect Sobol indices for turbine #14

	u	σ_u	α	θ
Mudline Mx DEL	0.33	1.15	0.00	0.15
Mudline My DEL	0.36	0.44	0.00	0.06
Tower base Mx DEL	0.48	1.14	0.00	0.20
Tower base My DEL	0.47	0.43	0.00	0.07
Tower top Mx DEL	0.09	0.59	0.00	0.05
Tower top My DEL	0.73	0.44	0.00	0.08
Tower top Mz DEL	0.10	0.58	0.00	0.04
Blade root Mx DEL	0.13	0.52	0.01	0.07
Blade root My DEL	0.66	0.38	0.01	0.04
Shaft Mz DEL	0.86	0.35	0.00	0.08
Rotor RPM	0.90	0.01	0.00	0.05
Blade pitch	0.95	0.00	0.00	0.01
Mean power	0.93	0.00	0.00	0.03
Mudline Mx Max	0.69	0.17	0.00	0.02
Mudline My Max	0.33	0.26	0.01	0.02
Tower base Mx Max	0.73	0.15	0.00	0.02
Tower base My Max	0.39	0.21	0.01	0.01
Tower top Mx Max	0.17	0.34	0.05	0.03
Tower top My Max	0.69	0.07	0.00	0.01
Tower top Mz Max	0.11	0.57	0.01	0.05
Blade root Mx Min	0.65	0.22	0.00	0.01
Blade root My Min	0.98	0.27	0.00	0.03
Shaft Mz Mean	1.22	0.06	0.00	0.05

The dependence of the turbine loads and power on wind speed is natural, since the wind turbines are built around the idea of adjusting their operation to extract maximum energy from the wind. To study further potential dependencies, we make an additional series of simulations where the wind speed is fixed, and Sobol indexes are computed for the remaining 3 random variables; i.e., conditional dependencies, per wind speed bin. An example of the outcome of these computations is shown in Figure 19 for the mean power output (left plot), and the blade root flapwise bending fatigue (right plot). The plots indicate that the influence of wind direction is highest around 10m/s, which can be explained with the wind turbine thrust coefficient that is typically highest around such wind speeds – and subsequently the wake effects are strongest. Another interesting observation is that the effects of wind direction and turbulence seem to be inversely proportional to each other: larger effect of the wind direction (wakes) means the effect of turbulence is less significant, and vice versa. Further, the wind direction (and wake effects) are the most critical factor affecting the power production, however for the loads the turbulence is more significant. The effect of wind shear seems to be marginal for the majority of the considered cases, both what concerns Figure 19 and Table 4.

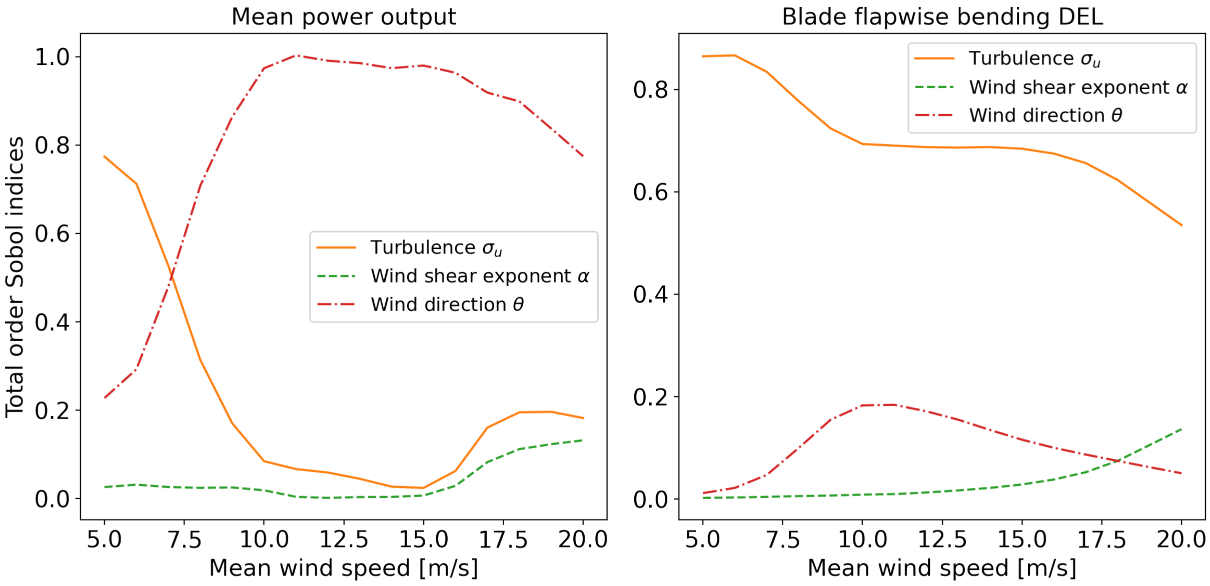


Figure 19 Total-effect Sobol indices as function of mean wind speed.

3.4. Comparison with SCADA measurements

Once all necessary surrogate models have been trained, power and loads prediction requires only four ambient inputs – wind speed, turbulence, wind shear and wind direction. If these are available from measurements, it is possible to compute power and loads snapshots and time histories for individual turbines as well as for the entire wind farm. For the Teesside wind farm, four years of SCADA data with 10-minutes resolution are available for all turbines, as well as wind speed, turbulence and wind shear from a nearby mast. Load measurements are not available, however the power measurements at least can be used to verify the predicted wake deficits – both for the reference Hawc2 simulations and for the surrogate model-based toolchains.

The met mast is located relatively far from the wind farm and is partially exposed to different inflow conditions. We choose therefore to obtain as much of the ambient wind information as possible from the nacelle anemometer readings. These measurements are affected by wakes, and in order to get the ambient wind speed and turbulence we need to only consider turbines that are in the free wind. This requires taking the wind farm geometry into account, and to determine the wind direction first. We use the algorithm described in [17], which can be briefly described as follows:

1) For a given 10-minute period, compute a preliminary wind direction by taking the average of all nacelle yaw directions;

2) Using the wind farm layout, determine the wind turbines for which there are no any upwind turbines within a direction sector of $\pm 15deg$ from the preliminary wind direction. This group of turbines is considered to be in free wind conditions for this particular 10-minute period.

3) An estimate of the ambient wind conditions (wind speed, wind direction and turbulence) is made by averaging the measurements from the free-wind turbines which are under normal operation (status=1) for the given 10-minute period. Note that for the wind direction it is the circular average that needs to be evaluated.

Finally, the met mast data are used to verify the predictions for wind directions where the inflow at the metmast and at nearby turbines is similar. The wind shear exponent is also estimated based on the met mast data as it is not possible to estimate from wind turbine readings only (wind speed readings are only available at one location – the nacelle anemometer – for each turbine).

Prior to the assessment of environmental conditions, the SCADA data need to undergo a preparation procedure - filtering, transformations, identification of turbine operating status, and similar. Since data engineering is not the primary focus of this report, we choose to omit the data preparation part. Similarly, we choose not to publish the resulting distributions of environmental conditions, as for the comparison we are focusing on a single wind turbine – T14, all wind directions and a few wind speeds – but without using any considerations about the probability distributions of these factors.

Figure 20, Figure 21, and Figure 22 show a comparison of the power output predictions at T14, for three wind speeds: 9m/s, 5m/s, and 11m/s, respectively.

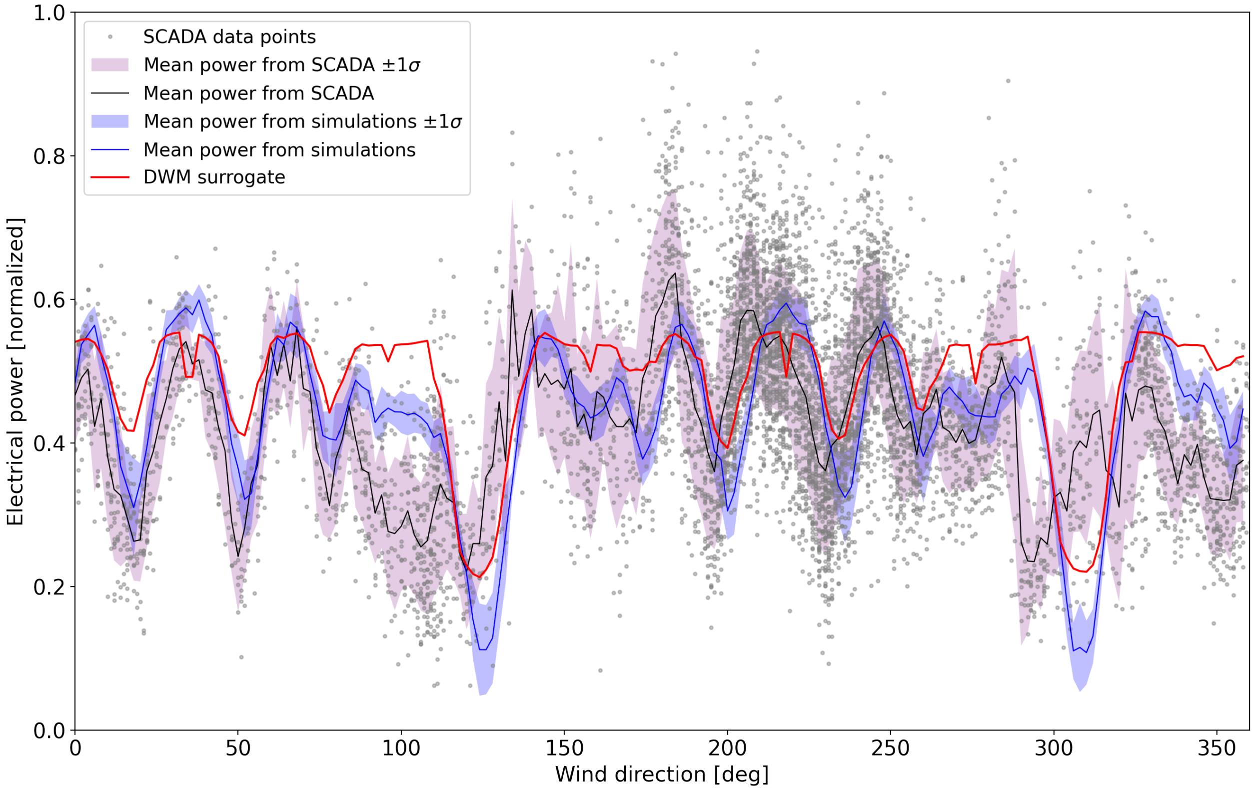


Figure 20 Comparison of model predictions of power output to SCADA measurements, at 9 m/s wind speed

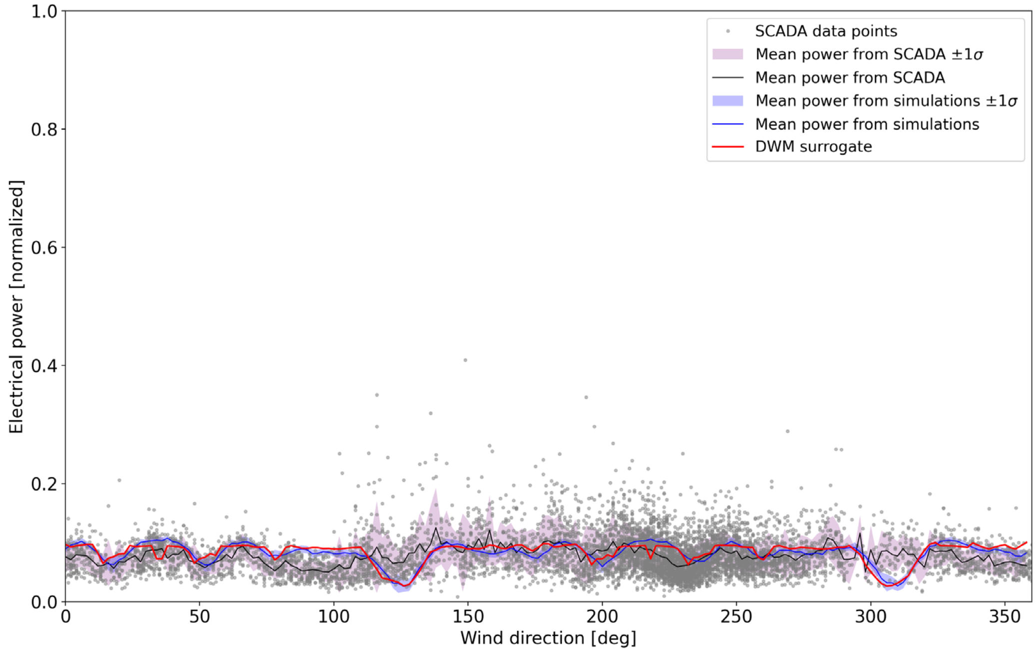


Figure 21 Comparison of model predictions of power output to SCADA measurements, at 5m/s wind speed

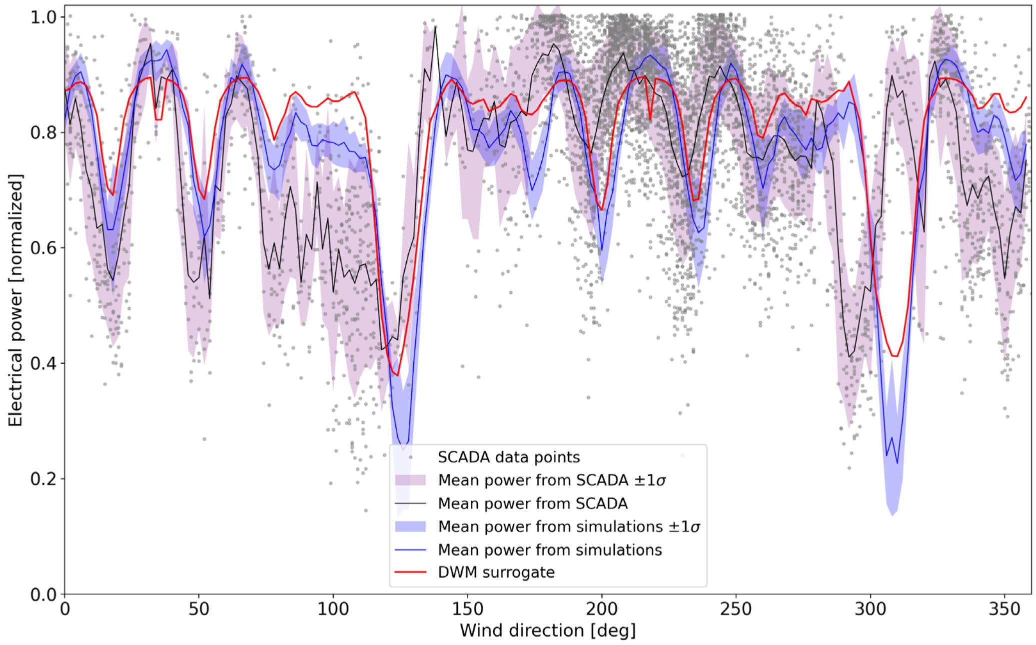


Figure 22 Comparison of model predictions of power output to SCADA measurements, at 11m/s wind speed

The main conclusions from the comparison are:

- The wake deficits predicted with Hawc2 simulations with the DWM model have in general a good agreement with the SCADA data. There is some mismatch of the deficit locations at wind directions above 180°, a potential explanation is uncertainty in the wind direction estimation;
- The surrogate model shows the same general behavior and captures the dependence on wind speed as well as the wake-free conditions, but performs less well for the deepest deficits. The

latter was already seen in earlier validations against simulation data and is most likely due to the uncertainty in load simulations that require extra amount of data to train surrogates;

- There are some specific situations where the models can underpredict or overpredict the power output. Some of these situations, especially the power overprediction, are likely due to measurement uncertainty on the wind speed and other ambient conditions from the SCADA data.
- The scatter in the simulations is less than that in the SCADA data. This is expected as in the simulations all environmental inputs are precisely defined, while in the SCADA data there are potentially random factors that are unknown or at least not measured but that have an influence on the power output, and which cannot be captured by the models.

References in Section 3

[1] Dimitrov, N. K. (2019). Surrogate models for parameterized representation of wake-induced loads in wind farms. Wind Energy, 22(10), 1371–1389. https://doi.org/10.1002/we.2362

[2] Yan, C., Pan, Y., & Archer, C. L. (2019). A general method to estimate wind farm power using artificial neural networks. Wind Energy, 22(11), 1421–1432. https://doi.org/10.1002/we.2379

[3] Dimitrov, N., & Natarajan, A. (2021). Wind farm set point optimization with surrogate models for load and power output targets. Journal of Physics: Conference Series, 2018(1), 012013. https://doi.org/10.1088/1742-6596/2018/1/012013

[4] Javier De Jaureguizar Pla (2021) Effect of wake model uncertainties on power output estimates for wind farms with wake steering control. DTU Wind Energy-M-0439

[5] Mads M. Pedersen, Paul van der Laan, Mikkel Friis-Møller, Jennifer Rinker and Pierre-Elouan Réthoré. (2019, February 12). DTUWindEnergy/PyWake: PyWake (Version v1.0.10). Zenodo.

[6] Zong, H., & Porté-Agel, F. (2020). A momentum-conserving wake superposition method for wind farm power prediction. Journal of Fluid Mechanics, 889.

[7] Crespo, A., & Herna, J. (1996). Turbulence characteristics in wind-turbine wakes. Journal of wind engineering and industrial aerodynamics, 61(1), 71-85.

[8] Katic, I., Højstrup, J., & Jensen, N. O. (1987). A Simple Model for Cluster Efficiency. In W. Palz, & E. Sesto (Eds.), EWEC'86. Proceedings. Vol. 1 (pp. 407-410).

[9] J. H. Halton, “On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals,” Numerische Mathematik, vol. 2, no. 1, pp. 84–90, 1960.

[10] Dimitrov, N. K., Kelly, M. C., Vignaroli, A., & Berg, J. (2018). From wind to loads: wind turbine site-specific load estimation with surrogate models trained on high-fidelity load databases. Wind Energy Science, 3(2), 767–790. https://doi.org/10.5194/wes-3-767-2018

[11] Larsen, T. J., & Hansen, A. M. (2007). How 2 HAWC2, the user's manual. Risø National Laboratory.

[12] Kingma, D. P. and Ba, J. (2014) Adam: A method for stochastic optimization. https://doi.org/10.48550/arXiv.1412.6980

[13] Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. Reliability Engineering and System Safety, 93(7), 964–979. https://doi.org/10.1016/j.ress.2007.04.002

[14] Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Mathematics and Computers in Simulation, 55(1-3), 271–280. https://doi.org/10.1016/S0378-4754(00)00270-6

[15] Dimitrov, N. K., Natarajan, A., & Kelly, M. C. (2015). Model of wind shear conditional on turbulence and its impact on wind turbine loads. Wind Energy, 18(11), 1917–1931. https://doi.org/10.1002/we.1797

[16] Kelly, M. C., Larsen, G. C., Dimitrov, N. K., & Natarajan, A. (2014). Probabilistic Meteorological Characterization for Turbine Loads. Journal of Physics: Conference Series (Online), 524(1), 012076. https://doi.org/10.1088/1742-6596/524/1/012076

[17] Dimitrov, N. K., & Natarajan, A. (2019). From SCADA to lifetime assessment and performance optimization: how to use models and machine learning to extract useful insights from limited data. Journal of Physics: Conference Series (Online), 1222(Conf. 1), 012032. https://doi.org/10.1088/1742-6596/1222/1/012032

4. Wind farm wake effect parameterization using surrogates for local wind estimates

4.1. Introduction and objective

The objective of the approach documented in this section is to elaborate an efficient turbine specific wake induced wind field parameterization and to achieve surrogate modelling for loads prediction (extreme and fatigue loads prediction). The overall goal is to build a general computational framework that could be used in an arbitrary wind farm layout.

The wake field hitting a rotor has some specificities. Indeed, when a turbine is located inside a warm, it experiences a modified wind field compared to the ambient one. Typically, the wake field exhibits a reduced mean wind speed and an increased turbulence intensity.

The current design approach for wake induced load analysis just consists in using a kind of equivalent turbulence intensity that would cause the same damage as the wake induced wind field.

For that purpose, IEC standard [1] proposes the concept of effective (or equivalent) turbulence intensity based on the work of Frandsen [2]. The effective turbulence is an imaginary one that would produce the same damage as the real ambient turbulence distribution. This concept may easily be extended to take account the combined effect of ambient turbulence and wake added turbulence.

This effective turbulence intensity has a major practical advantage: in one value it encompasses the whole combined effects of ambient turbulence distribution, wind direction distribution and wake added effects. This advantage is of great importance in structural load analysis because it allows us to save a lot of computation time. Nevertheless, the drawback is that the effective turbulence intensity is dependent on a material parameter (in particular the Wöhler exponent m).

The expression for effective turbulence intensity relies on a major assumption: the damage equivalent load is proportional to the wind speed standard deviation, or equivalently to the standard deviation of the random part of the structural response. This assumption implies that deterministic component of the structural response is somewhat overwhelmed by the stochastic component (no wake mean deficit).

With this assumption, and following [2], the effective turbulence intensity (conditioned on mean wind speed) that combines ambient turbulence and wake added turbulence may be expressed as:

$$I_{eff}^{m} \cong \left(1 - \sum_{j=1}^{N} p_{w,j} \left(\frac{I_{0,j}}{I_m}\right)^{m}\right) I_m^m + \sum_{j=1}^{N} p_{w,j} I_{T,j}^{m} \qquad \text{Eq. 1}$$

Where, $I_m$is the ambient effective turbulence intensity, $I_{0,j}$ is the directional ambient effective turbulence intensity, $I_{T,j}^{m}$ is wake effective turbulence intensity and $p_{w,j}$ is a weighting factor.

The expression for $p_{w,j}$ is:

$$p_{w,j} = b_j 2\theta_{w,j} f_{wd,j} \qquad \text{Eq. 2}$$

Where:

$$b_j = \frac{\sqrt{\pi}}{2}\frac{\sum_{i=1}^{m}\binom{m}{i}\frac{\alpha_j^i}{\sqrt{i}}}{\sum_{i=1}^{m}\binom{m}{i}\alpha_j^i} \qquad \text{Eq. 3}$$

and $\theta_{w,j}$ is the characteristic width of wake numb. j, $f_{wd,j}$ is the value of the PDF of wind direction around wake center and α_j^i is a wake added turbulence coefficient.

But this simplified approach does not account for the specific shape of the wake velocity field and may lead to inappropriate (conservative or under-conservative) load predictions especially for rotors operating in partial-wake conditions (see Figure 23).

The objective is hence to implement an efficient « computing route » for wake induced extreme and fatigue load prediction for arbitrary layout with current research engineering software that should tackle:

- ✓ added turbulence and
- ✓ velocity deficit shape.

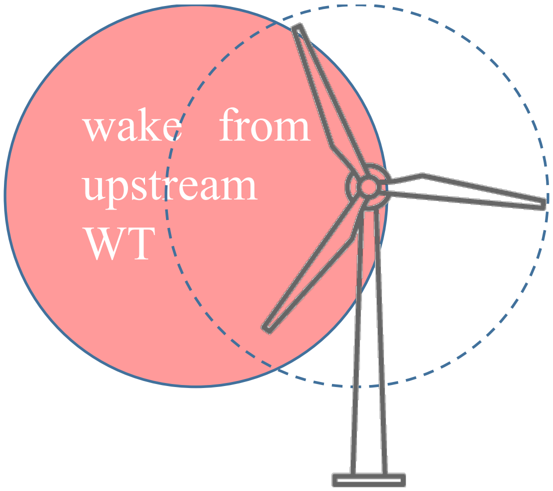


Figure 23 Partial wake condition: the blade may experience here an additional load variation due to the partial overlap of the wake. This phenomenon is not explicitly taken into account in the current design standards [1]

One of the main constraints of the present study is that this parameterization must be compatible with current IFPEN research engineering software (or model chain) for static wake analysis and it should also stay reasonable in terms of number of parameters, especially if the model chain involves fully coupled FEA software. This limitation is desirable for two reasons: [i] some surrogate modelling techniques are limited in terms of number of dimensions (for Kriging, the inversion of the covariance matrix is expensive) [ii] it avoids the curse of dimensionality in terms of design of experiment (DOE) when considering fully coupled FEA simulation cost.

The wind farm parametrization proposed by IFPEN is an alternative to the one proposed by DTU. It considers a "mean wind speed profile-oriented parametrization". Basically, it consists in finding a suitable parametrization of the yz vertical plane distribution of the time average wind velocity field (from FarmShadow^{TM}[2] or any other wake sources) at the wind turbine location with an appropriate model function. The objective is finally to elaborate a mapping between the quantities of interest (here fatigue and extreme load) as function of the parameters that describe the wind within a farm. This mapping is provided here by the use of a surrogate model based on a Gaussian process (or Kriging). As the parametrization proposed here is only dependent on the wind field profile hitting a wind turbine, it does not depend on the wind farm layout.

Besides, this parametrization relies on the assumption that the numerical tool to compute wake induced mean wind speed in a farm is a static analysis tool (like FarmShadow^{TM} for example) and hence do not solve any partial differential equations: it can be the case if it implements wake analytical models such as Bastankhah and Porté-Agel Gaussian wake model [5]. As a consequence, there is no particular benefits to build a metamodel of such numerical tools as the CPU time dedicated to a single wake computation is quite negligible in comparison with the CPU time required by a fully coupled aero-servo-hydro-elastic computation.

After a brief description of the model function and the associated "wake induced parameters" needed to fit a yz vertical plane time average wind velocity field inside the farm, the fitting process is presented. Then, to obtain a sort of wind field profiles "catalog" adjusted to the Teesside wind farm, a design of experiments (DoE) for FarmShadow^{TM} simulations is processed. Based on this catalog and the fitting process, the joint distributions of the wake induced parameters can be used to build a relevant DoE for DeepLinesWind^{TM} simulations [6]. Those simulations are fed with modified turbulent wind fields obtained with in-house scripts in association with a turbulence box generator. These simulations serve then as training data for Gaussian process surrogate models in order to map the wake induced parameters and some quantities of interest (Damage Equivalent Load or extreme load). Ultimately, with such a surrogate model, DEL or extreme load may be instantly requested for a specific set of wake parameters and Global Sensitivity Analysis (GSA) may also be conducted as it is illustrated at the end of this section.

4.2. Parametrization approach

4.2.1. Description of the model function

Let us define the y-axis as the local transverse axis of the rotor and the z axis as the vertical axis[3]. The figure below displays a typical time average wind velocity vertical profile that could hit the rotor (the rotor contour is represented with a red circle in Figure 24). This figure clearly exhibits a velocity deficit perturbation zone due to the presence of one or several upwind turbines. The objective of the parametrization is to find the best set of wake induced parameters that fit a model function to the mean wind velocity profile.

---

[2] FarmShadow^{TM} is a static wake analysis software developed by IFPEN and dedicated to the modelling of wind farm flows based on engineering wake models [3], [4]

[3] The z vertical axis is indifferently the local or the global axis

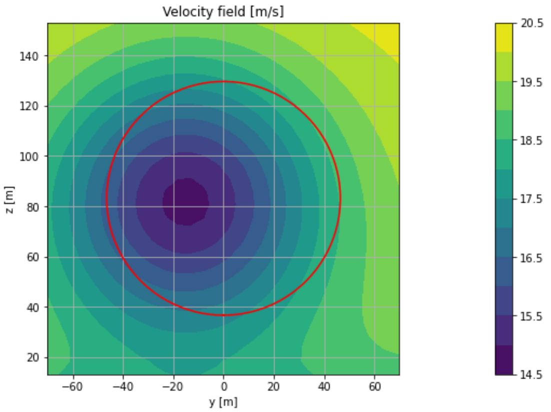


Figure 24 An example of a mean longitudinal wind velocity profile hitting a rotor (the red circle represents the border of the rotor)

The problem statement is the following: we are looking for a function $\varphi(y,z)$ to fit a mean wind velocity field profile given by a static wake analysis software like FarmShadow^{TM} or any other sources (LES or wind measured wind sources for instance) with basic nonlinear least squares regression.

To address this problem, we suggest choosing a function based on some physical considerations. The model function used is a radial basis function (RBF) made of two-dimensional Gaussian functions that are combined using linear summation rule (see Eq. 4). This model function has been motivated by the Bastankhah and Porté-Agel Gaussian wake model [5] although this model is not entirely satisfactory in near-wake situations [3].

$$\varphi(y,z) = u_0 \left(\frac{z}{\hat{z}}\right)^{\alpha} \left(1 - \sum_{i=1}^{N} a_i e^{-\frac{1}{\sigma_i^2}\left[(y-\hat{y}_i)^2 + z^2\right]}\right) \quad \text{Eq. 4}$$

Where u_0 is the free stream wind speed at hub height $\hat{z}$, a_i is the amplitude of the ith wake, $\hat{y}_i$ is the y-coordinate of the ith wake center, σ_i is the characteristic wake width of the ith Gaussian, α is the wind shear exponent and N the total number of Gaussians. This number is not a priori fixed but will depend on the complexity of the wind field profile. The amplitude $a_i$may be viewed as the maximum velocity deficit due to the presence of an upwind turbine. Note that we assume here that the vertical location of the center of the wake is equal to the hub height. This assumption is justified by the fact that we limit the application of this parametrization in this section to the Teesside wind farm which is composed of turbines mounted on monopile foundations. It may however be necessary to relax this assumption for the case of a floating wind turbine like the South Brittany study case of HIPERWIND. The vertical position of the center of the wake may then vary according to the floater motions.

As said before, this model function is inspired by the Porté-Agel analytical model implemented in FarmShadow^{TM} and also by the summation rule. Indeed; for a (x,y,z) point in the farm, the wind velocity is written as:

$$u(x,y,z) = u_0(z) - \sum_{i=1}^{N} u_i^* \, \Delta u_i(x) \Psi_i(r(y,z))$$

Where $u_0(z) = u_0 \left(\frac{z}{\hat{z}}\right)^{\alpha}$ and u_i^* is the mean wind speed velocity of the i$^{\text{th}}$ upwind rotor, Ψ_i is the shape function for the wake, r is the radial distance from the wake center and $\Delta u_i$is the maximum value of the velocity deficit. If the velocity deficit model chosen is based on the Bastankhah and Porté-Agel Gaussian wake model for example, the shape function is a Gaussian function. Note also that the maximum value of the velocity deficit is dependent on the wake width σ and the value of the thrust coefficient C_T. The following expression relates the maximum velocity deficit with both the wake width and the thrust coefficient:

$$\frac{\Delta u}{u_0(z)} = \left(1 - \sqrt{1 - \frac{C_T}{8\left(\frac{\sigma}{D}\right)^2}}\right) \tag*{Eq. 5}$$

Where C_T is the thrust coefficient, σ is the characteristic wake width, and D is the rotor diameter.

Based on this expression, we observe that when C_T is increasing, the maximum velocity deficit increases. Note also that the wake width σ depends on the turbulence intensity. Indeed, the turbulence intensity is the most influential parameter that governs the wake growth rate.

With such a parametrization, 3 parameters are needed for each Gaussian function (a_i, σ_i and $\hat{y}_i$) and three additional "global" parameters I_u,(average turbulence intensity over the rotor) u_0 and α. As a consequence, 9 parameters are needed for two Gaussians and 6 parameters for one Gaussian. The number of Gaussians required will be determined during the fitting process (see §4.2.2). The "global parameters" may be viewed as ambient conditions parameters while the 3 parameters describing the shape of each Gaussian function may be viewed as wake induced parameters.

This set of parameters is the one necessary for the mean wind speed representation. Let us note here that the spatial distribution of turbulence intensity hitting a wind turbine is ignored and the influence of upwind rotors on the turbulence is only taken into account with the single value of the turbulence intensity averaged on the rotor only.

The quality of the fitting will be evaluated with the classical metric:

$$RMSE = \sqrt{\sum_{i} \frac{\left(vel(y,z) - \varphi(y,z)\right)^2}{N}} \tag*{Eq. 6}$$

Where N is the number of points comprised in the yz wind field plane and $vel(y,z)$ is the actual wind field value at each (y,z) points.

The RMSE value will be normalized by the standard deviation of the free stream wind speed at hub:

$$NRMSE = \frac{RMSE}{\sigma_{u_0}} \qquad \text{Eq. 7}$$

Where σ_{u_0} is the free stream wind speed standard deviation.

4.2.2. Wind field plane fitting methodology

Starting from a yz wind field plane (or wind field vertical map) and a model function, the objective here is to find the set of parameters such that the model function described above best fits the wind profile. Note that if the 3 ambient parameters associated to the yz wind field plane are already known, it is not necessary to include them in the fitting process: in that case, only the wake-induced parameters are concerned. The non-linear least square fitting process is performed with the Levenberg-Marquardt algorithm through the SciPy library. The Jacobian matrix of the model function is numerically evaluated but could be also easily analytically computed. The initial guess is determined with differential evolution genetic algorithm to avoid local minimums. During the fitting process, if the targeted error criteria is not achieved, additional Gaussian functions may be added if and only if this additional function significantly reduces the normalized RMSE. The addition of Gaussians is hence subjected to a kind of relative performance criteria expressed in terms or relative error change:

$$\frac{RMSE_i - RMSE_{i+1}}{RMSE_i} > PC$$

Where $RMSE_i$ is the root mean squared error calculated for the ith Gaussian and PC is the performance criteria.

This performance criteria has been introduced in order to prevent, as far as possible, the addition of unrealistic Gaussians: it is also a way to ensure a compromise between accuracy and computational performance as the addition of a Gaussian significantly increases the number of wake parameters. In addition, the fitting process is constrained by upper and lower bound for each parameter. The following tables give examples of these definition domains for the case of Teesside wind farm for which the rotor diameter D is set to 93 m.

	Lower bound	Upper bound
$\hat{y}_1$	-160 m	+160 m
σ_1	20	120
a_1	0	0.7

Table 5 Lower and upper bounds for parameters in the case of one Gaussian for Teesside case

	Lower bound	Upper bound
$\hat{y}_1$	-160 m	0 m
$\hat{y}_2$	0 m	+160 m
σ_i	20	120
a_i	0	0.7

Table 6 Lower and upper bounds for parameters in the case of two Gaussians for Teesside case

When a Gaussian function i is added, the range of the location parameter $\hat{y}_i$ is limited to a portion of the total horizontal range (see Table 6). More precisely, for N Gaussians, the location parameter $\hat{y}_i$ is comprised in the range [-160+320i/N] for i=1 to N. Note that the yz wind field plane for each rotor is hub centered and covers the whole rotor disk area.

As a quick illustration, let us consider a simple application. The wind farm studied here is the Teesside wind farm which consists of 27x2.3 MW turbines located off the coast of north east England in the North Sea (see Figure 25). All turbines have a hub height of 83.1 m and a rotor diameter of 93 m. This wind farm exhibits an average value of separation distance between turbines (considering the 3 nearest neighbor turbines only) of 4.7 D, where D is the rotor diameter. The minimum separation distance is 3.6 D.

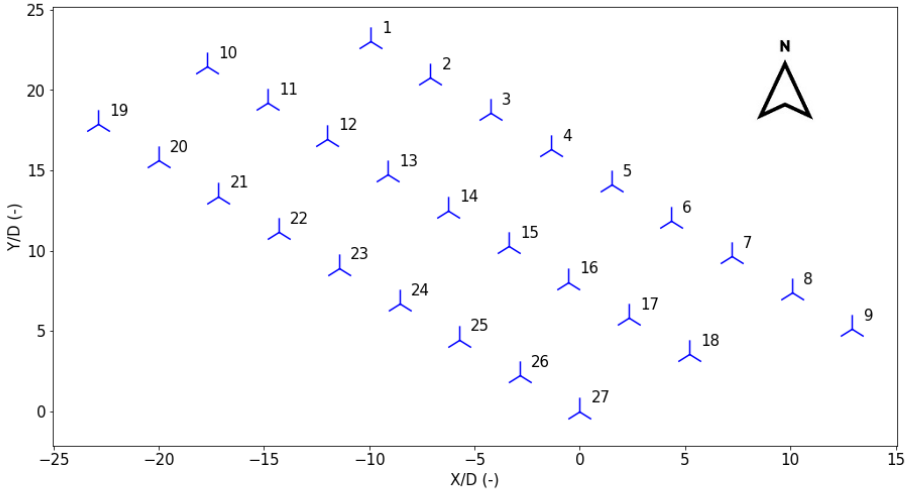


Figure 25 Teesside wind farm (X and Y designate the global axis coordinates)

FarmShadow^{TM} offers several options for wind velocity deficit and added turbulence modeling. The choice was made to focus on the most advanced analytical wake models. The velocity deficit model chosen is the super Gaussian deficit [3], and the wake added turbulence model is that of Qian-Ishihara [7]. To account for interaction between the turbines and their wakes, linear superposition is assumed. The relationship between the wind speed and the height is modeled with a wind profile power law (shear exponent α). The thrust coefficient C_T vs U curve used for FarmShadow^{TM} computations is plotted below.

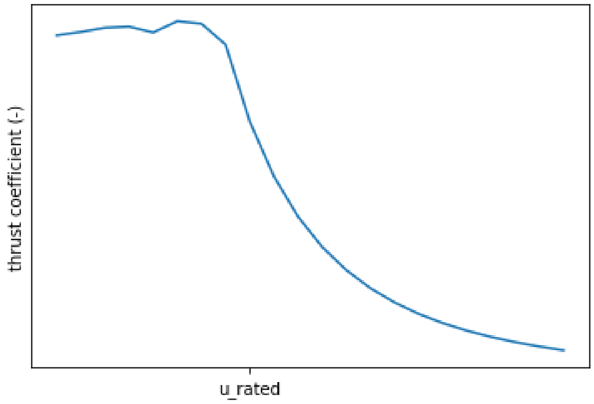


Figure 26 Thrust coefficient C_T vs wind speed

This curve has been obtained with doing constant wind simulations with DeepLineWind^{TM} for steady wind and different step of wind speed (from 4 m/s to 25 m/s).

In the subsequent plots, the wind speed is normalized wrt. to the free stream wind speed at hub height.

Figure 27 shows an example of FarmShadow^{TM} output for the Teesside wind farm with a wind coming from the west: the free stream wind speed u_0 is set to 9 m/s, the free stream turbulence intensity I_{u0} is set to 10 % the and wind shear power law coefficient α is set to 0.17.

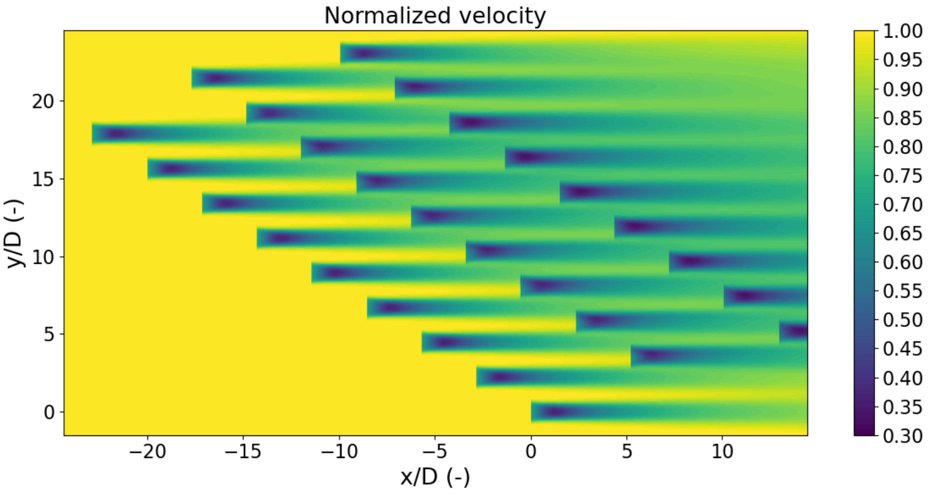
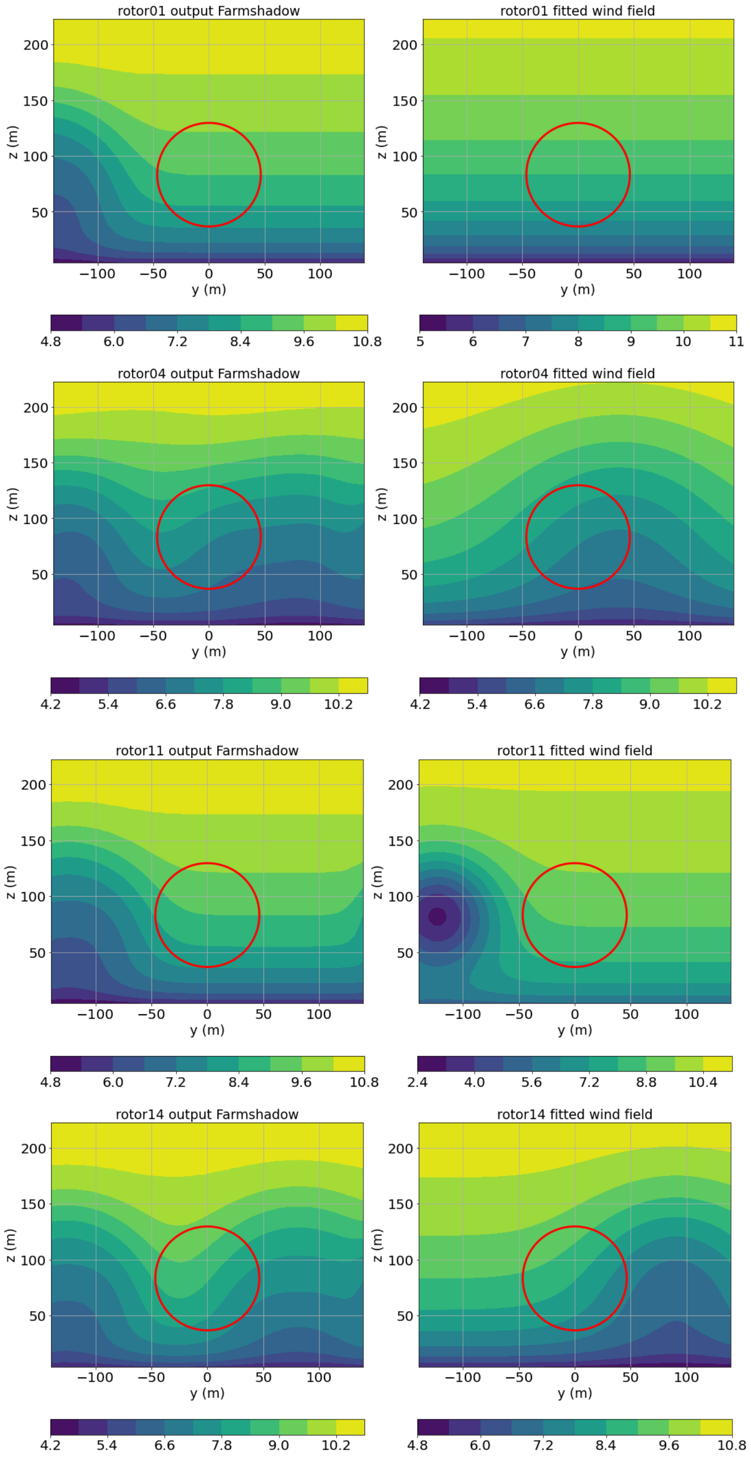


Figure 27 Hub-height normalized velocity at Teesside wind farm for a wind coming from west (u_0=9m/s, I_{u0}=0.1, α=0.17). The velocity is normalized with respect to the free stream velocity at hub height and D is the rotor diameter

We apply the fitting process to all the yz wind field planes evaluated at each rotor of the Teesside wind farm. The following figures only display some representative turbines. The rotor#23 is not affected by any wakes because this turbine relies on the front line. The rotor #14 is a rotor that belongs to the second line and that is affected by two wakes while the rotor #11, although belonging to the second line is affected by only one wake. Finally, the rotor #4 and the rotor #1 belong to the third line and are affected respectively by at least two wakes and one wake.


rotor01 output Farmshadow
rotor01 fitted wind field
rotor04 output Farmshadow
rotor04 fitted wind field
rotor11 output Farmshadow
rotor11 fitted wind field
rotor14 output Farmshadow
rotor14 fitted wind field
z (m)
y (m)
200
150
100
50
−100
−50
0
50
100
4.8
6.0
7.2
8.4
9.6
10.8
5
6
7
8
9
10
11
4.2
5.4
6.6
7.8
9.0
10.2
2.4
4.0
5.6
7.2
8.8
10.4

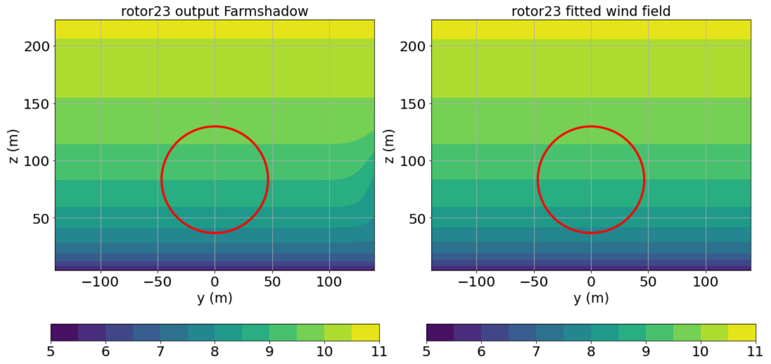


Figure 28 yz wind field distributions for rotor#1, #4, #11, #14 and #23. Pictures on the left are the outputs from FarmShadow$^{\text{TM}}$, pictures on the right are the fitted profiles

Looking at the wind farm scale it is clear that for a wind coming from west, the first wind turbine row (from rotor #19 to rotor #27) is not affected by an upwind turbine. Consequently, for this specific row, the yz wind field profile may be correctly represented by the usual three wind field characteristics namely u_0, I_{u0} and α (ambient conditions parameters). The second and the third row are mainly affected by the presence of one upwind turbine (see right of Figure 29).

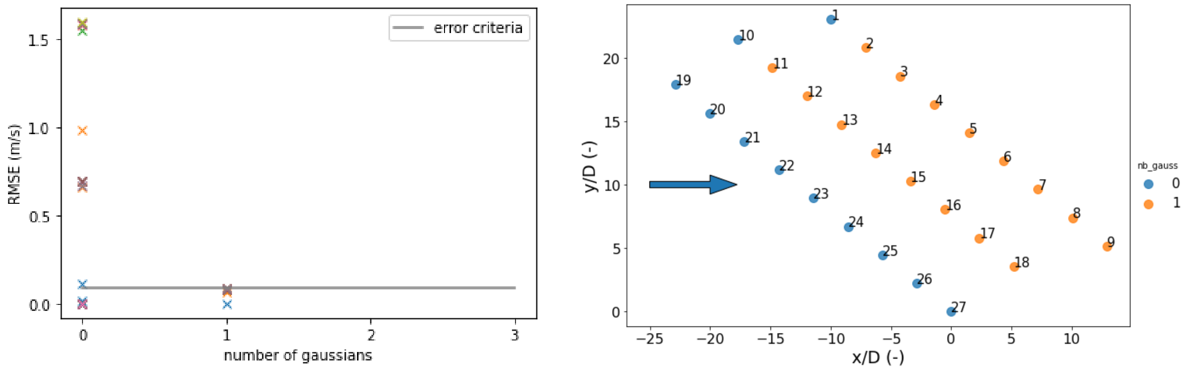


Figure 29 convergence curve of the fitting error (left) and wind farm colored by the number of Gaussians required (right)

In this example, the error criterion in terms of RMSE is equal to 0.09 m/s as we choose here to limit the fitting error (see Eq. 4) to 10 % of the wind speed standard deviation ($NRMSE < 0.1$). In this case, two Gaussians are enough to satisfy the error criterion or the relative performance criteria expressed here in terms of a relative error change > 33%. (left of Figure 29)

4.3. Design of experiments

4.3.1. FarmShadowTM DoE

The previous example has focused on a specific ambient condition. In order to cover various environmental conditions, a design of experiments has to be created. Building a FarmShadow^{TM} design of experiments serves two purposes:

1. the first purpose is to determine the maximum number of Gaussians required to achieve good approximation of "typical" wind field profiles without compromising the surrogate modeling objective because of the curse of dimensionality issue. For that purpose, a sort of "catalog" of wind profiles at turbine should be established from an appropriate FarmShadow^{TM} Design Of Experiment of Teesside wind farm,
2. the second purpose is to identify parameters ranges and distributions for each case (0, 1 or 2 Gaussian case).

It is important to precise that FarmShadow^{TM} is a numerical tool that does not solve any partial differential equations: it implements analytical wake models only. Therefore, there is no particular benefits to build a metamodel of FarmShadow^{TM} as the CPU time dedicated to a single FarmShadow^{TM} computation is quite negligible.

This design of experiments for FarmShadow^{TM} is here reduced to a 4-dimensional hypercube because, for a specific wind farm, only the ambient conditions are necessary. The ambient wind characteristics required are u_0, I_{u0}, α and wind direction. With the aim of optimizing the hypercube filling, 999 points have been generated using space-filling Latin Hypercube Sampling (LHS) within the followings bounds and with uniform distributions except for wind speed (truncated Weibull) as it is considered as an independent variable with the Rosenblatt transformation. The bounds for the other variables are conditionally dependent on wind speed. The conditional distributions and variable ranges for wind speed, turbulence, and wind shear used in the present study are based on the same procedure proposed by Dimitrov in [8] and adapted here to the Teesside wind farm case.

	Lower bound	Upper bound
u_0 (m/s)	4	25
θ (°)	0	360
σ_{u_0} (m/s)	$0.025\mathrm{u}_0$	$0.16(0.012\mathrm{u}_0{}^2 + 0.75\mathrm{u}_0 + 5.6)$
α (-)	$0.15 + 0.1(D/\hat{z})(u_{max}/\mathrm{u}_0)$	$0.05 - 0.05\left[1 - 0.4\ln(0.5\,D/\hat{z})^2\right](u_{max}/\mathrm{u}_0)$

Table 7 Lower and upper bounds for the ambient wind characteristics

Where D is the rotor diameter, $\hat{z}$ is the hub height, α is the shear exponent, and u_{max} is the upper bound of the wind speed (or cutout wind speed, that is 25 m/s). Finally, σ_{u_0} is the standard deviation of the wind speed velocity.

Wind direction is uniformly distributed in [0°,360°]. The distribution parameters of the truncated Weibull for u_0 are A=15 and k=2. These values differ from the site-specific conditions but are chosen to fill the space better. A wind direction equal to 0° refers to a wind coming from the west and it rotates anticlockwise.

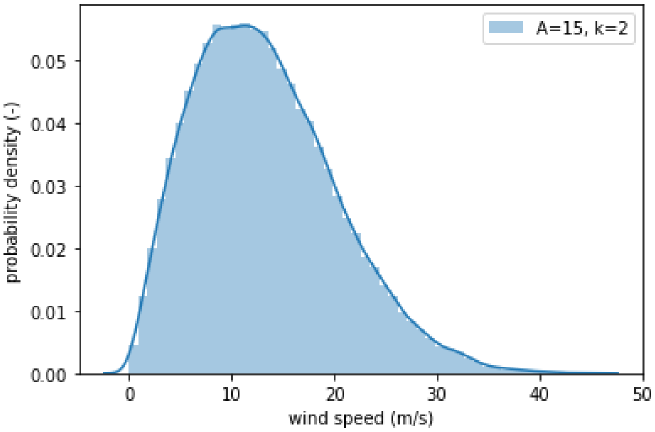


Figure 30 Weibull distributions for ambient wind speed u_0.

The space-filling LHS algorithm used is the one available in the R library MOLHD [9] that, according to the library documentation, *"searches the optimal Maximin design using columnwise exchange algorithm coupled with the simulated annealing algorithm and several computational shortcuts to improve efficiency"*. More information about this LHS that maximizes the minimum distance among the design points may be found in [10]. The use of such a space filling LHS ensures a better distribution of the sample points and hence a faster convergence of the sample statistics.

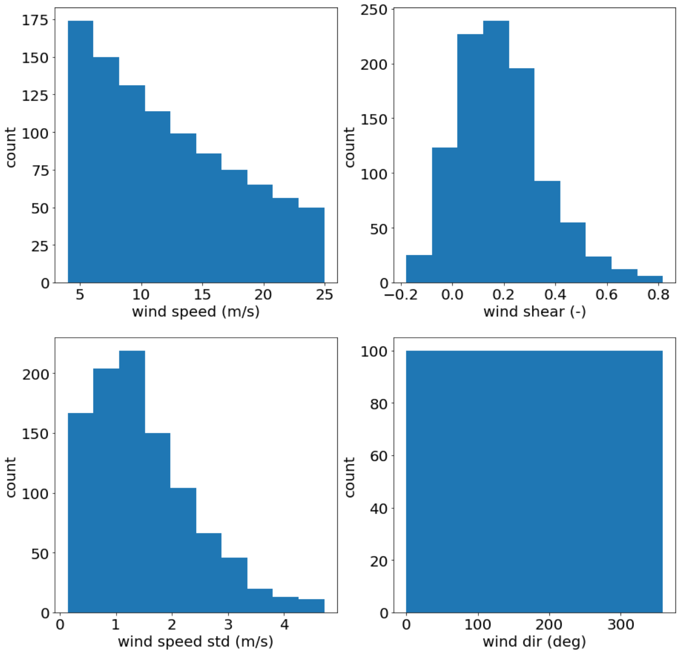


Figure 31 LHS samples histograms for the 4 ambient conditions parameters

As expected, the maximin LHS offers a quite low discrepancy compared to a random sampling. The following figure display the 1-dimensional scatter of the wind direction (reduced here to the 150°-210°

sector for the sake of clarity). The discrepancy may be viewed as a measure of both the highest and lowest density of a sequence. A high value of discrepancy means that there is some area with high density and other with low density. A low discrepancy on the contrary means that the density is quite uniform (an equidistributed sequence has a discrepancy equal to 0).

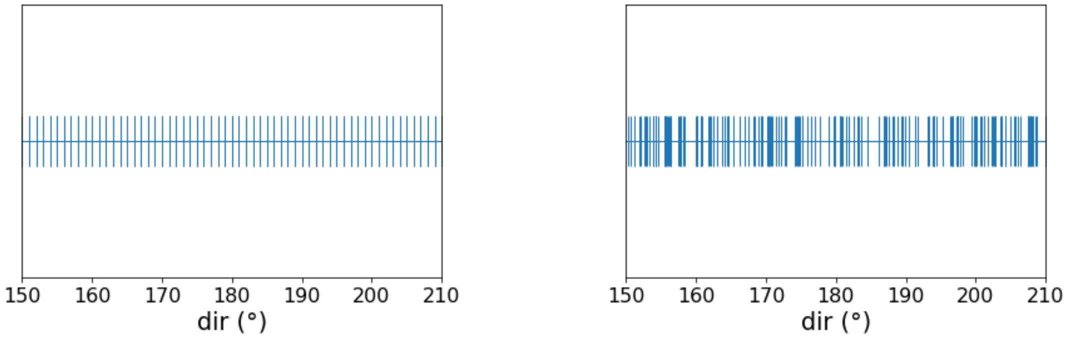


Figure 32 1-dimensional scatter of the space filling LHS (left) and random sampling (right). Only the wind direction parameter is displayed as it is the only parameter that is unconditionally uniformly distributed

The following figures show the scatter plots for the 999 samples space-filling LHS.

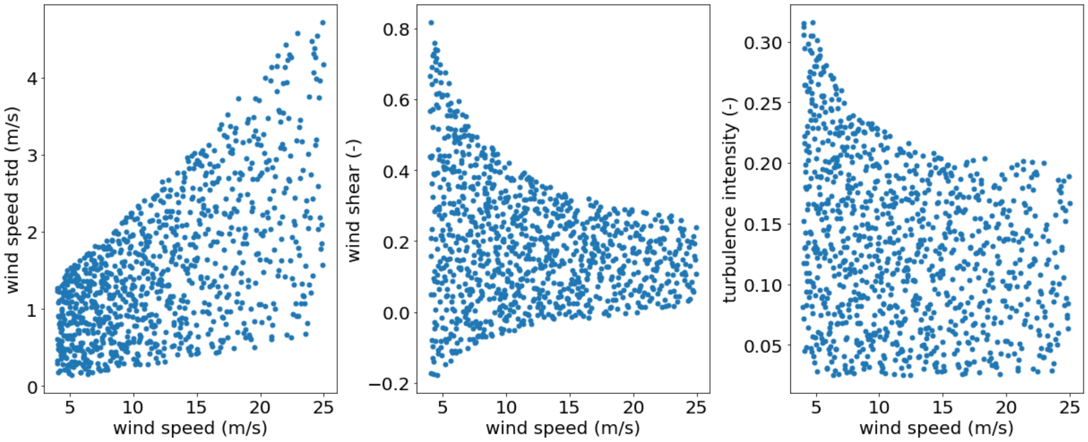


Figure 33 2-dimensional scatter of the space filling LHS

4.3.2. Fitting process

Based on the optimization process previously described (see §4.2.2), the whole 26 973 wind field profiles are processed with the fitting methodology described above.

At the end of the fitting process, we get 26 973 set of parameters. Depending of the complexity of the wake, these set of parameters may be composed of 3 parameters (no Gaussian, ambient case), 6 parameters (one Gaussian that is one major wake hitting the rotor), 9 parameters (two Gaussians). Note that for the Teesside case study, the 3 parameters of the no Gaussian case (I_{u0}, u_0 and α) are known and are hence not considered in the fitting process.

It is also interesting to note that, for this Teesside case study, 44 % of all the wind profiles are correctly fitted with the ambient wind field (no wake), 43 % with one Gaussian and only 2.5 % with two Gaussians. No wind profiles exhibited a level of spatial complexity that would have required a third Gaussian function.

The following histograms informs us of the proportion of “complex” wind profiles with respect to ambient profiles assuming that the number of Gaussians required by the fitting process is correlated to the wake complexity. The proportion of complex wind profiles is higher for wind speed values below 10 m/s and decreases from 10 m/s to 25 m/s (see Figure 34). This tendency may be explained by the shape of the thrust coefficient C_T vs U curve. When U is above 10 m/s, C_T decreases and then the maximum velocity deficit decreases. Below 10 m/s, C_T is high and almost constant. Consequently, the maximum velocity is also high and almost constant (see Eq. 5). The perturbation zone, that may be defined as the portion of the yz wind field where the velocity deficit is above a certain threshold, increases with the maximum velocity deficit, raising the risk of its overlapping with the rotor disk.

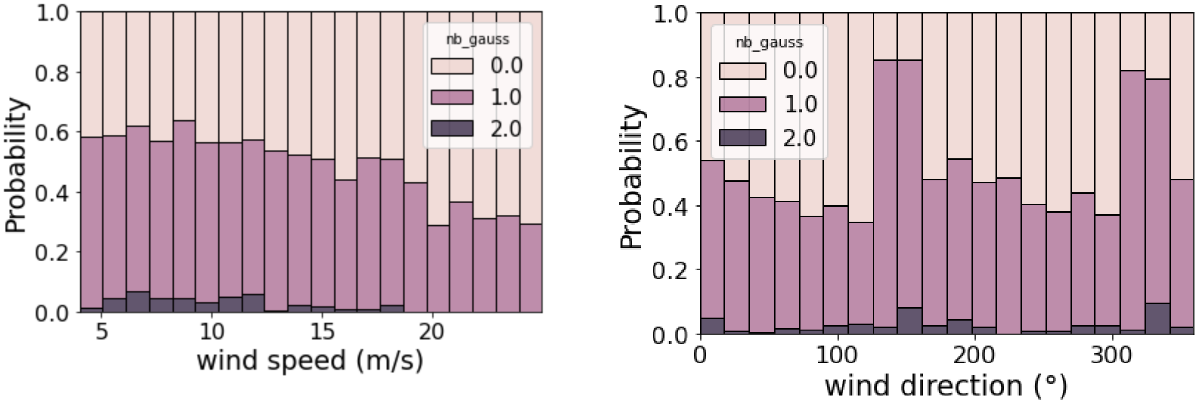


Figure 34 Wake complexity proportion as a function of mean wind speed (top left) and wind direction (top right). The proportion of complex wakes (1 Gaussian or more) is higher when C_T is high and for 142°/320° wind direction.

For the sake of clarity, let us consider the velocity deficit behind an arbitrary turbine. Figure 12 represents the evolution of the velocity deficit value at $y=\pm D/2$ as a function of the mean wind speed (see Figure 34). This function looks similar to the one observed for C_T vs u except for wind speed below 10 m/s. In this range, C_T is constant, but the velocity deficit increases with u. This is due to the sharp drop of the turbulence intensity (see Figure 33) that tends to increase the maximum velocity deficit value.

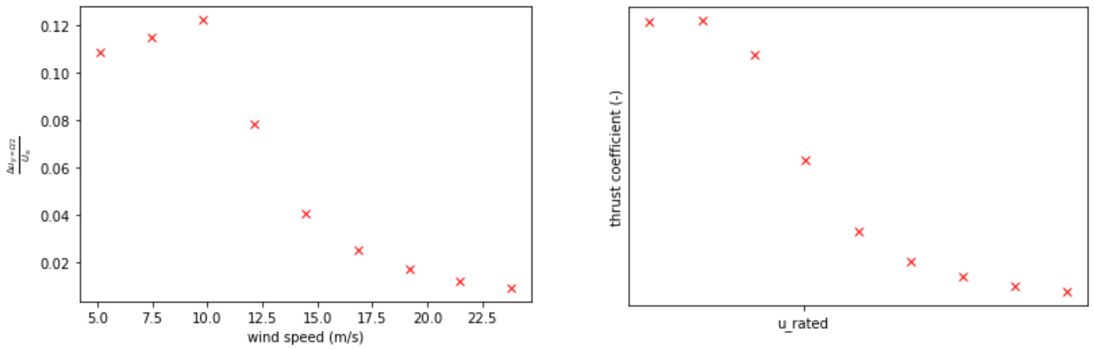


Figure 35 Velocity deficit at y=±D/2 wrt. mean wind speed value (left) and thrust coefficient evolution with mean wind speed value (right)

Regarding the wind direction, it is also interesting to note that one particular direction may be emphasized on the wake complexity given by Figure 34. This direction (142°/322°) is the one corresponding to the case where the wind is aligned with the turbine rows (see *Figure 36*) and is the one corresponding to the higher proportion of complex wind profiles. However, one should also note that one gaussian is enough to represent accurately the local wind profile on any wind turbine for this specific wind direction. We can also note that two gaussians are the maximum needed for any configuration of wind direction and turbine of the Teesside layout. It can then be expected that a limited number of parameters can represent the local wind on any wind turbine independently of the wind farm size. This is a strong difference with the alternative approach presented in Section 3, for which the parameters encompass all upwind turbines.

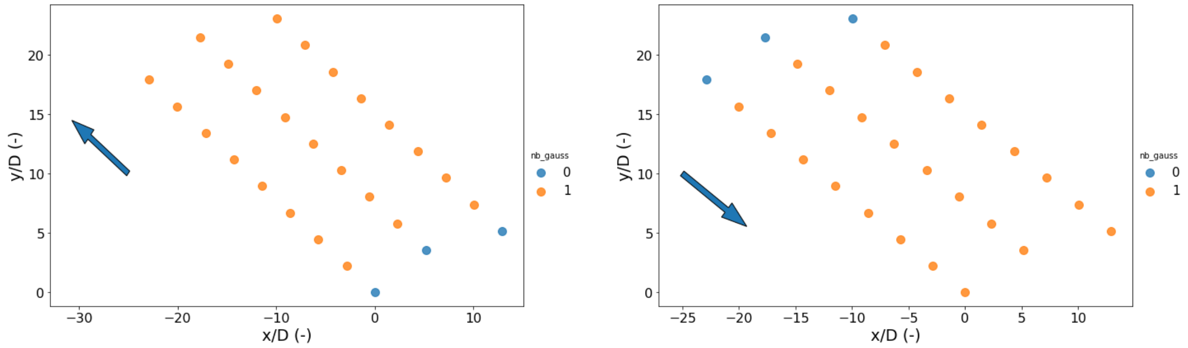


Figure 36 Proportion of wake impacted turbines for 142°/320° wind direction. In this direction almost 90% of the turbines are affected by a wake-induced wind field

Out of a total of 26 973 wind profiles, almost 87 % and 95 % exhibit a normalized RMSE below 10 % and 20 % respectively (see *Figure 38*). The mean value if NRMSE is 5 % while the median is only 1.77 %. Those values have also to be confronted with regards to the wind direction. If we plot the mean value of NRMSE conditionally to the wind direction, we find that the maximum mean value occurs at 142° and 320°. Those directions are not the most probable direction for the Teesside wind farm for which the most probable wind direction is west (see *Figure 37*). If we had to compute a weighted damage considering the wind distributions, the impact of the high values of NRMSE would be alleviated.

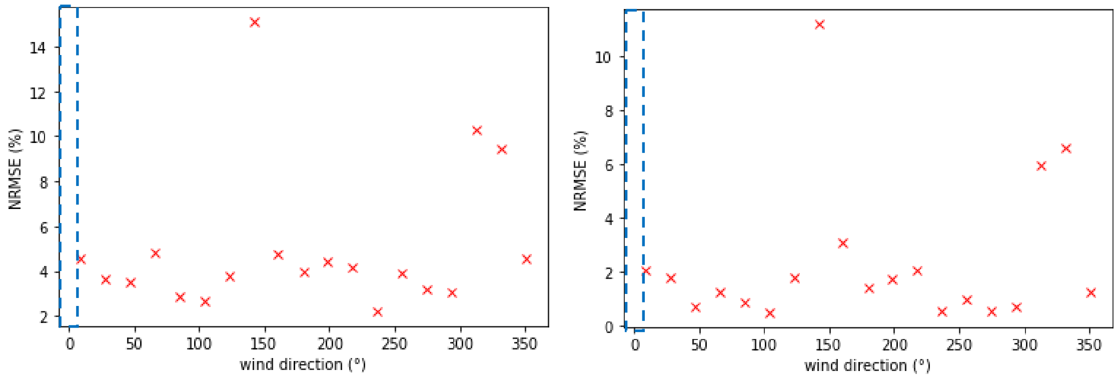


Figure 37 Mean (left) and median (right) NRMSE value conditionally to the wind direction. Most probable direction is indicated with a blue dashed rectangle

As noted before, the remaining profiles that exceed the error criteria target would not be better fitted with a third Gaussian as the performance criteria is not satisfied. Improvement would therefore require another summation rule or additional parameters for the model function. The latter could be the vertical location of the Gaussian function z_i, or a super-Gaussian parameter instead of n=2 as it was observed both numerically and experimentally, that the wake velocity profiles are not purely Gaussian [3]. Finally, a kind of covariance matrix $\Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$ could be introduced in the model function that could differentiate y and z direction in terms of wake dimensions.

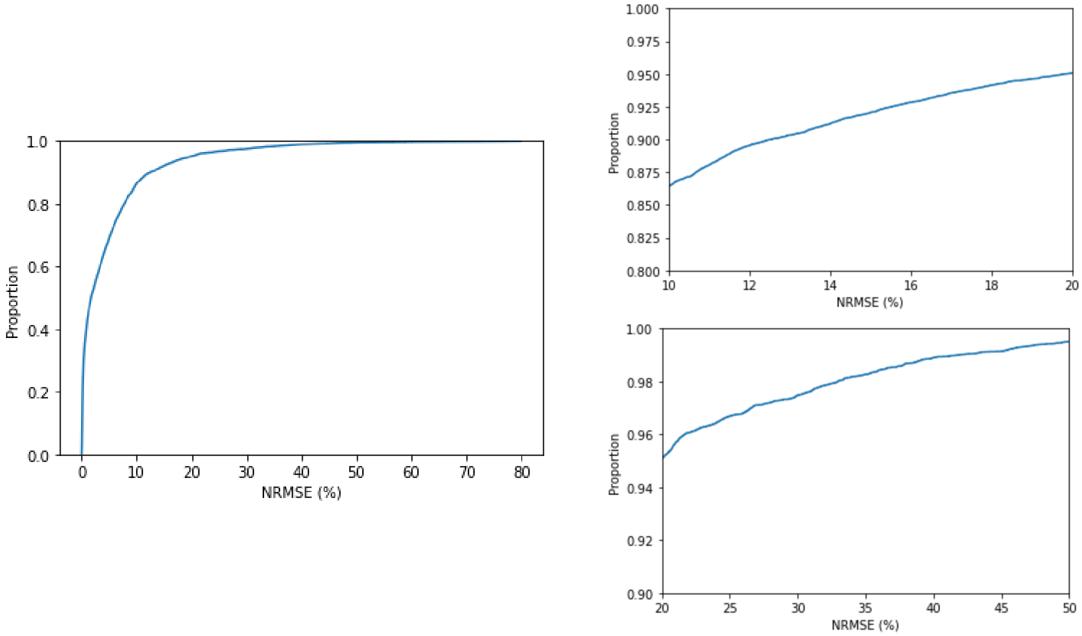


Figure 38 Normalized cumulative frequency (empirical cdf) of the normalized RMSE

It is interesting to observe the statistical characteristics of the fitted wake parameters. The following table (see Table 8) summarizes the main statistical quantities for the 1 Gaussian case and the 2 Gaussians case.

The histograms of wake parameters are represented in Figure 39 and Figure 40 for the 1 Gaussian case and the 2 Gaussian cases respectively.

For the 1 Gaussian case, the mean value of the amplitude (or maximum velocity deficit) is 0.2 with a relatively high dispersion. In addition; the location parameter exhibits a bimodal distribution for the 1 Gaussian case that could reflect here a combined effect of the borders of the horizontal axis of the yz mean wind velocity profile and the elliptical shape of the wake due to the superposition of more than one Gaussian. In that case, the center of the model function stays almost fixed while a sufficient portion of the wake has not crossed the border of the yz mean wind velocity plane. The elliptical shape could be represented for example by a two-dimensional Gaussian and a covariance matrix.

Note also some accumulations at the bounds both for the wake width parameter and for the location parameter (see Figure 39 and Figure 40). These accumulations may reflect a suboptimality of the fitting process that should be solved in order to improve the overall accuracy of the parametrization strategy proposed in this section. A more appropriate model function could be a way to solve this issue.

	1 gauss.		
	a_1	$\hat{y}_1$	σ_1
mean	0.2	-0.9	54.8
std	0.2	51.1	25.8
min	0.0	-160.0	21.6
25%	0.1	-42.3	36.9
50%	0.2	-2.3	45.9
75%	0.3	40.6	64.3
max	0.7	160.0	120.0

	2-gauss.					
	a_1	$\hat{y}_1$	σ_1	a_2	$\hat{y}_2$	σ_2
mean	0.3	-97.2	61.1	0.3	90.6	45.9
std	0.2	34.7	29.8	0.2	22.6	15.4
min	0.0	-160.0	21.3	0.0	33.7	21.3
25%	0.2	-115.0	36.4	0.2	74.4	36.6
50%	0.2	-94.4	48.7	0.3	89.3	41.4
75%	0.4	-71.7	80.6	0.5	102.7	51.0
max	0.7	-11.1	120.0	0.7	160.0	117.5

Table 8 Statistics of the wake parameters for 1 and 2 Gaussian cases

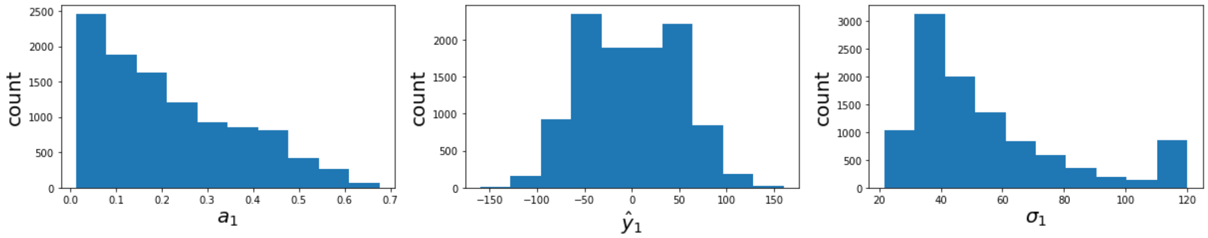


Figure 39 Histograms of the 3 wake induced parameters for the 1 Gaussian case. A bimodal distribution is clearly visible for the location parameter y_1

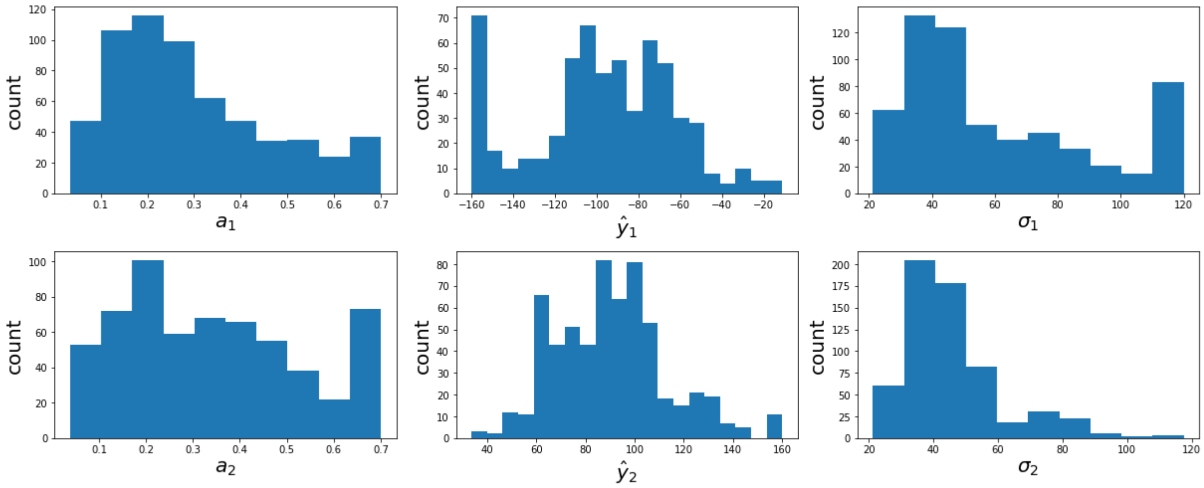


Figure 40 Histograms of the 6 wake induced parameters for the 2 Gaussian case. Note the presence of accumulation points for the location parameters at the bounds of the domain (-160 m)

With the previous distribution, a second Design Of Experiments will be established for performing DeeplinesWind^{TM} (DLW) simulations.

4.3.3. Design of experiments in terms of parameters

The main purpose of FarmShadow^{TM} simulations in this section is to identify and restrict the range of the wake induced parameters for each case (0, 1 or 2 Gaussian case) and have some understandings of their distributions. As we must build a design of experiments in terms of parameters, the goal here is obviously to limit the number of DeepLinesWind^{TM} simulations to appropriate combinations of wake induced parameters. In this sense, the design of experiments is farm dependent. Nonetheless, if we have no restriction on the number of simulations, the limitation imposed on the wake induced parameters combinations could be removed. Notwithstanding this farm-based restriction, the parametrization strategy and the surrogate models are completely farm independent (it does not depend on the wind farm layout). Nonetheless, it is dependent on the wind turbine characteristics.

For each of the case (0, 1 or 2 Gaussian case), a specific DoE has to be elaborated in order to train a surrogate model. The design of experiments in terms of parameters is based on the distributions found with the fitting process of the 26 973 yz profiles described in the previous section. This DoE will be used for training a surrogate model, that is an approximation model devoted to substitute a more complex and time-consuming simulator. An ideal DoE is a one that exhibits both space filling properties and that preserves the joint distributions of the input variables. Here, for the sake of simplicity, we choose to combine of two types of DoE, the advantages of one counterbalancing the disadvantages of the second and inversely. Let us precise though, that this purely empirical approach has not been rigorously evaluated in terms of surrogate model performance:

- ✓ the first one is a Latin Hypercube Sampling space filling design built to stratify the marginal distributions found in the fitting process (see §4.3.2). The space filling Latin Hypercube Sampling (LHS) algorithm used is the one available in the R library MOLHD [9]. The drawback of this approach is that the sampling is unable to take into account the correlations between input variables. As a consequence, the Iman & Conover distribution-free approach [6] is used so as to include a correlation matrix into the LHS sampling. This method has the advantage to preserve the marginal distributions,

- ✓ the second DoE is a “full Monte-Carlo” (full MC) design of experiments (or simple random sampling) that has been built through a subsampling of 26 973 fitted profiles. This DOE is used because it has good projective properties and provides the advantage to handle complex input variable systems exhibiting a high level of correlations among explanatory variables as the joint distributions are preserved. However, this DoE has very bad space filling properties.

It should be noted that before sampling the parameter distributions, a multivariate outlier detection based on Mahalanobis distance is implemented. Each DoE contains 999 sample points. Consequently, for each case (0,1 or 2 Gaussian case) the total number of sample points is equal to 1998.

In the same way as for the parameters describing the mean wind speed, the sampling of rotor averaged turbulence intensity is based on the distribution coming from FarmShadow$^{\mathrm{TM}}$ outputs.

4.4. DLW simulations

4.4.1. Model description

Fully-coupled aero-servo-hydro-elastic simulations are performed in DeepLinesWind™ The DLW model of the Teesside turbine is a standard finite-element one with beam elements. It is implementing all the stiffnesses in the blade and even blade torsion. A classical representation of the airfoils with polar data is used for BEM aerodynamics computation. Specific prebend and shear center data were missing but elastic and aerodynamic center are accurately input. The hydrodynamics loads are computed with a Morison formulation. As for the soil-structure interaction, the degrees of freedom of translation along the z axis and of rotation around the z axis are blocked at seabed. All 4 other degrees of freedom are purely diagonal stiffness terms and modeled by flexjoints or springs. In DLW, only the stiffness proportional damping is modeled whereas data is provided for mass and stiffness damping. Thus, a value for the stiffness proportional damping is computed for this model to match the value of damping in the blade and tower for their first modes respectively. The DTU controller is designed for pitch-regulated variable speed wind turbines. It is built for Hawc2 [11] but a wrapper is provided that is used by DeepLinesWind™.

4.4.2. Wind modifications

Assuming a set of parameters $(a_i, \hat{y}_i, \sigma_i)$, a 3900 s turbulent wind field has to be generated accounting for the spatial variation of mean wind speed through a model function (see §4.2.1 and §4.2.2). This generation is held by the modification of a reference turbulent wind field generated with a turbulence box generator (TurbSim [12] for example). This modification is operated with the help of specific in-house python scripts and consists in offsetting the mean wind speed at each point of the grid specified in the TurbSim input file. The offset value is computed with the value of the model function and the set of wake parameters. A 28x28 point square grid of three-component wind vector is used. Let us recall that the spatial distribution of turbulence intensity is ignored and the influence of upwind rotors on the turbulence is only taken into account with a unique value turbulence intensity value averaged over the rotor.

In TurbSim, for the generation of a turbulent wind field with a target turbulence intensity I_u and a target mean wind speed $\bar{u}$, the wind field at each (y,z) points and for each time t is generated with respect to this expression:

$$u(t) = \tilde{u}(t,y,z) + \bar{u}$$

Where $\tilde{u}$ is the fluctuating part of the wind speed (zero-mean stochastic process) and $\bar{u}$ is the mean wind speed value.

The offsetting consists in replacing the fluctuating part in the following manner:

$$\tilde{u}^*(t,y,z) = \tilde{u}(t,y,z) - \bar{u} + \varphi(y,z)$$

To be sure to obtain the turbulence intensity averaged on the rotor after modification, the target I_u requested in the TurbSim .inp file is the turbulence intensity averaged on the rotor disk and the mean wind speed requested is the harmonic mean of φ over the rotor disk.

Below, an example of turbulent wind field (mean wind speed u_0=19.86 m/s, ti=5.8%) modified by a single Gaussian wake (a_1=0.25, y_1=-15 m and σ_1=37.8 m).

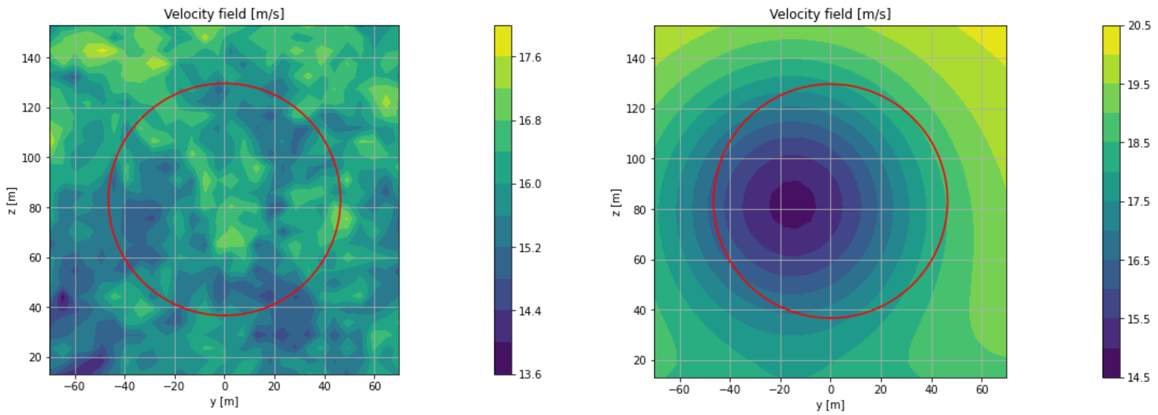


Figure 41 yz turbulent wind field plane from TurbSim (left) is modified by a single Gaussian wake (right)

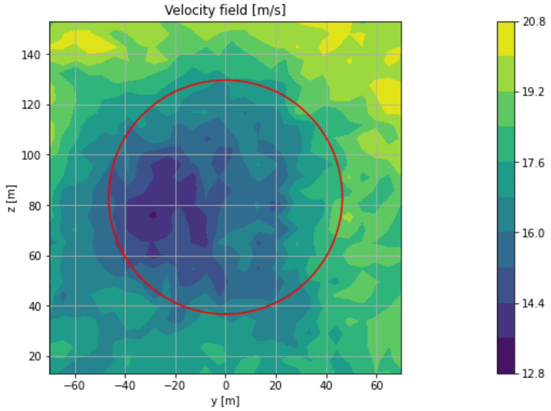


Figure 42 yz turbulent wind field plane with the wake modification

4.4.3. Turbulence spectra parameters

The spectral velocity tensor is here described by the Kaimal model for which the integral scale parameters L_k are defined in [1]. Moreover, the IEC exponential spatial coherence is used and the default spatial coherence parameters are chosen (see [12] for more details).

$$Coh_{i,j_K} = exp\left(-a_k\sqrt{\left(\frac{fr}{u_{hub}}\right)^2 + (b_K r)^2}\right)$$

Where f is the frequency, r the distance between points i and j, a_k and b_k are the input coherence decrement and offset parameters and u_{hub} is the mean wind speed at hub.

With the aim of having a useful part of the time series equal to 3600 s, 3900 s simulations are executed. Hence, we expect to reach the steady state after 300 s. These aeroelastic load simulations are run for a single turbulent wind time series (one seed namely one turbulence time histories). Let recall that a turbulent wind may differ with the value of the stochastic seed as a turbulent wind is a random realization of a stochastic process. The length of the useful part of the simulation is here 3600 s as recommended by IEC standard [1].

4.4.4. Output computation (DEL and extreme loads)

The quantities of interest are the fatigue damage equivalent load (DEL) at different location of the wind turbine and the extreme load. The 1Hz DEL is calculated using the following expression:

$$DEL = \sqrt[m]{\sum_{i=1}^{N} \frac{\Delta F_i^m N_i}{N_{ref}}}$$

With m is the the Wöhler exponent, N_{ref} the reference number of cycles (N_{ref}=3600 in case of 1 hour simulation) and ΔF_i is the load range cycles obtained with rainflow counting algorithm [13]. The Wöhler exponents are defined as m=3 for steel (this is the case for the tower) and m=10 for composite material (this is the case for the blades). No corrections is made to account for the mean load effect on the damage equivalent load.

The extreme load is simply computed as the global maximum value over the simulation length.

Two locations are of interest here: blade root and tower bottom. For each of these locations, the bending moment is evaluated along two directions: edgewise and flapwise for the blade root and fore-aft and side to side for the tower bottom.

4.5. Surrogate model and GSA

A metamodel is a response surface approximation. It is a simplified model, which once obtained, allows to avoid the calls to a costly simulator. Here the objective is to build a surrogate model that would provide the mapping between both the ambient condition parameters and the wake induce parameters on one side and the load (fatigue loads or extreme loads) on the other side.

The mapping of the wake parameters and the loads implies the use of a surrogate model. Here, a surrogate modeling has been built and trained with 80 % of all the simulation data available: this set is termed “train set”. The 20 % remaining data are used to evaluate the accuracy of the surrogate: this set is termed “test set”.

The simulation data are the output computed at the end of the DeepLinesWind^{TM} simulations (see §4.4). Note that the simulation data are filtered beforehand so as to remove outliers and idling cases. The idling cases are identified when the blade pitch position reaches 90°.This is done with the multivariate outlier detection procedure already described (see §4.3.3).

	Wake complexity								
	0 gauss.			1 gauss.			2 gauss.		
DoE	initial	Idling	outlier	initial	Idling	outlier	initial	Idling	outlier
DoE 1	999	118	278	999	0	106	999	0	109
DoE 2	999	110	123	999	7	146	999	0	173
Total	1369			1739			1716		

Table 9 DLW simulations data

Finally, before training, the input data are scaled (or normalized in the [0,1] interval so as to cope with the variation of the ranges of the inputs (or features). The output is not normalized but is just scaled by a factor of 1000 (thus from kN to MN).

4.5.1. Kriging

Many types of surrogate modeling are available but kriging regression (or Gaussian process regression [14]) has been selected because it can interpolate the outputs and it produces both predictions and associated confidence intervals and it is suited for low dimension problems (here the number of dimensions is at most equal to 9 in the two Gaussian case).

Kriging is a Bayesian prediction method which aims at building a function $\mathcal{M}$, assuming it is a realization of a Gaussian process, in which we have a finite number of observing points: $\chi = \left\{\boldsymbol{x}^{(i)}, i = 1, \ldots., m\right\}$.

Finally, we can predict under certain hypotheses, the values taken by the simulator like a Gaussian random variable: $\hat{Y}(\boldsymbol{x}) = \left[\, \mathcal{M}(\boldsymbol{x}) |\, y_i = \mathcal{M}\left(\boldsymbol{x}^{(i)}\right), i = 1, \dots, m\right] \sim \boldsymbol{\mathcal{N}}(\mu_{\hat{Y}}(\boldsymbol{x}), \sigma^2_{\hat{Y}}(\boldsymbol{x}))$ where the mean $\mu_{\hat{Y}}$ and the variance $\sigma^2_{\hat{Y}}$ are analytically known.

The interpolation is of interest, but the main interest of the prediction allows us an associated probability value to the prediction. A 1D illustration is displayed on Figure 43 with the function $f(x) = x\ sin(x)$.

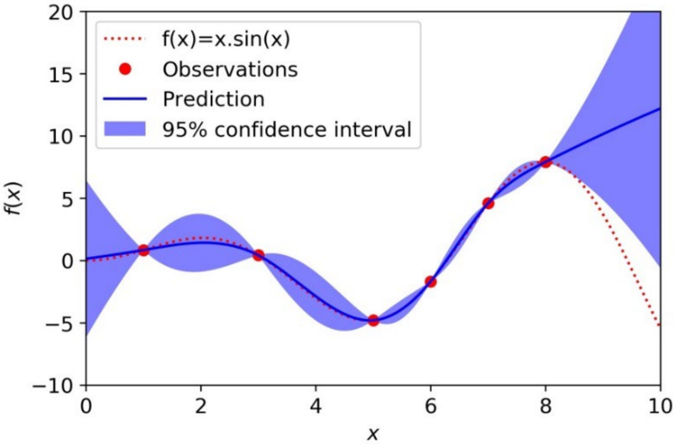


Figure 43: Interpolation property of the best linear unbiased predictor for one-dimensional function: $f(x) = x\ sin(x)$ (from *[15]*))

A Gaussian process is hence entirely specified by its mean and its covariance function. A square exponential function is chosen here for the covariance and an anisotropic kernel is used. Kriging hyper-parameters are optimized by maximizing the log-marginal likelihood with a quasi-Newton algorithm and within specified bounds (L-BGFS-B algorithm). This optimization is done with the train set.

4.5.2. Kriging performance

The NRMSE (normalized root mean squared error) is used as a measure of the error made on the test set in order to facilitate the comparisons between QOI. Different kind of normalization are here proposed: normalization by the mean value, by the standard deviation value and by the amplitude. It should be noted that the amplitude normalization is said to be more sensitive to the sample size while the mean normalization may lead to some difficulties in the interpretation when the quantities of interest (QoI) have large differences in their mean value. An interesting illustration is the comparison of the surrogate performances for the blade root edgewise DEL and blade root flapwise DEL. The respective mean NRMSE exhibit significant differences while the std RMSE is of the same order. This is due to the fact that the blade root edgewise DEL has a higher mean value than flapwise but a lower dispersion as this load is mainly driven by the weight load. Hence, NRMSE values must be carefully compared.

RMSE Normalization		
mean	std	maxmin
$\frac{\sqrt{\sum_i \frac{(y_{test}-\widehat{y_{test}})^2}{N}}}{\overline{y_{test}}}$	$\frac{\sqrt{\sum_i \frac{(y_{test}-\widehat{y_{test}})^2}{N}}}{\sigma(y_{test})}$	$\frac{\sqrt{\sum_i \frac{(y_{test}-\widehat{y_{test}})^2}{N}}}{\max(y_{test})-\min(y_{test})}$

Figure 44 Three kind of RMSE normalization

Where $\widehat{y_{test}}$ is the predicted values on test set, y_{test} is the actual values in the test set and N is the sample size.

One surrogate model is trained for each quantity of interest and for each wake complexity case (0, 1 or 2 Gaussian case). Indeed, the dimension of the input space is dependent on the wake complexity. Then, for each [input,output] couple, a specific surrogate model is needed, that is 3 (wake complexity case) x 8 (QoI) = 24 surrogate models. The training process is operated with scikit-learn python library.

To evaluate the performance of the surrogate models, the RMSE for each QoI and each kind of inputs is compared. The quantities of interest are summarized below:

	Fatigue loads	Extreme loads
Blade root edgewise[4]	DEL_B_edge	Max_B_edge
Blade root flapwise	DEL_B_flap	Max_B_flap
Tower bottom fore-aft	DEL_TB_my	Max_TB_my
Tower bottom side-side	DEL_TB_mx	Max_TB_mx

Table 10 Combinations of inputs/outputs

	0 Gaussian case			**1 Gaussian case**			**2 Gaussian case**		
	mean	std	maxmin	mean	std	maxmin	mean	std	maxmin
Max_B_edge	3.05	22.58	5.49	3.25	21.61	5.28	2.49	17.81	4.44
Max_B_flap	3.42	19.46	4.04	3.26	12.7	2.96	3.84	13.6	3.53
Max_TB_mx	6.38	14.97	3.41	10.82	27.07	5.1	11.04	28.49	6.52
Max_TB_my	5.44	23	5.2	4.6	14.88	3.67	4.37	12.72	3.73
DEL_B_edge	0.44	9.96	2.1	0.74	11.84	2.54	0.56	10.21	2.37
DEL_B_flap	3.34	9.53	2.06	5.14	14.44	2.8	5.76	15.95	3.45
DEL_TB_mx	6.61	10.48	2.35	15.36	27.04	4.94	16.46	30.26	6.35
DEL_TB_my	3.82	8.24	1.99	7.74	19.5	3.77	7.06	19.55	3.87

Table 11 NRMSE (%) values for each QoI and each kind of inputs

It can be noticed that the blade root edgewise bending moment DEL prediction exhibits the lower valuer of NRMSE. This is probably due to the fact that this QOI is much less affected by the wakes as the edgewise

---

[4] Edgewise and flapwise bending moments are evaluated in blade root local axis that take account for the blade pitch angle

bending moment is mainly attributed to the weight load. Note also the large values of NRMSE for TB_mx that may suggest a more complicated wind speed dependency (see more details §4.5.3) Apart from this QOI, the other one exhibit comparable values of NRMSE. Lower values of NRMSE may be obtained with a larger amount of training data.

4.5.3. Global Sensitivity Analysis (GSA)

The sensitivity of each input parameter is assessed with the use of the surrogate model and the Fourier amplitude sensitivity testing which is a more efficient variance-based global sensitivity analysis than classical Monte-Carlo variance-based GSA. The Fourier Amplitude Sensitivity Testing (FAST) [16] method implemented in the OpenTURNS [17] Python package is used to evaluate the Sobol indices. Thus, the index $S_{i1...is}$ represents the proportion of the variance of Y due to the interaction between the variables $U_{i1},\cdots,U_{is}$. The first-order index considers the influence of U_i on Var(Y) without accounting for its interactions with the other input parameters. The total-order Sobol index S_{Ti} has been proposed in [18] to accumulate the effects of all the combinations of input parameters involving U_i. The sensitivity analysis is here conducted with uniform distributions for all inputs.

	a_1	$\hat{y}_1$	σ_1	a_2	$\hat{y}_2$	σ_2	u_0	I_u	α
Max_B_edge	0.004	0.019	0.011	0.053	0.074	0.025	0.686	0.265	0.002
Max_B_flap	0.012	0.024	0.022	0.015	0.034	0.015	0.840	0.105	0.005
Max_TB_mx	0.053	0.050	0.063	0.003	0.020	0.005	0.658	0.317	0.017
Max_TB_my	0.022	0.044	0.024	0.022	0.049	0.011	0.832	0.093	0.000
DEL_B_edge	0.009	0.026	0.013	0.095	0.111	0.030	0.701	0.126	0.013
DEL_B_flap	0.022	0.016	0.054	0.002	0.011	0.006	0.574	0.364	0.033
DEL_TB_mx	0.052	0.047	0.052	0.052	0.045	0.038	0.534	0.526	0.023
DEL_TB_my	0.039	0.044	0.021	0.027	0.028	0.009	0.337	0.670	0.013

Table 12 Total order Sobol indices for each QoI and each input – 2 Gaussian case

	a_1	$\hat{y}_1$	σ_1	u_0	I_u	α
Max_B_edge	0.067	0.211	0.027	0.585	0.379	0.002
Max_B_flap	0.082	0.142	0.025	0.718	0.221	0.005
Max_TB_mx	0.076	0.085	0.005	0.682	0.264	0.015
Max_TB_my	0.095	0.150	0.022	0.723	0.234	0.001
DEL_B_edge	0.089	0.358	0.025	0.565	0.206	0.010
DEL_B_flap	0.118	0.092	0.016	0.580	0.330	0.036
DEL_TB_mx	0.119	0.109	0.010	0.511	0.485	0.008
DEL_TB_my	0.069	0.062	0.010	0.318	0.670	0.009

Table 13 Total order Sobol indices for each QoI and each input – 1 Gaussian case

	u_0	I_{u0}	α
Max_B_edge	0.380	0.703	0.002
Max_B_flap	0.502	0.551	0.012
Max_TB_mx	0.789	0.261	0.013
Max_TB_my	0.512	0.595	0.001
DEL_B_edge	0.419	0.710	0.014
DEL_B_flap	0.598	0.370	0.082
DEL_TB_mx	0.705	0.388	0.001
DEL_TB_my	0.446	0.635	0.001

Table 14 Total order Sobol indices for each QoI and each input – 0 Gaussian case

Regardless of the wake complexity, the free stream mean wind speed at hub height u_0 and the rotor averaged turbulence intensity I_u are the most important factors that affect both the extreme loads and the fatigue loads. Surprisingly, wind shear exponent α does not seem to have a major impact on the loads: the impact is nonetheless a bit more pronounced for the loads experienced by the blades. For the 1 Gaussian case, note the significant influence of the wake location parameter regarding the blade root edgewise bending moment DEL.

4.6. Application case

The surrogate model prediction is tested with an application case. Let us consider the Teesside wind farm (see Figure 7) with a free stream wind speed u_0 = 9 m/s, free stream turbulence intensity I_{u0}= 10 % and wind shear α=0.17. We are interested here in predicting the loads with respect to the wind direction for the turbine #14 (see Figure 45) which is the worst turbine in terms of wake complexity based on several observations.

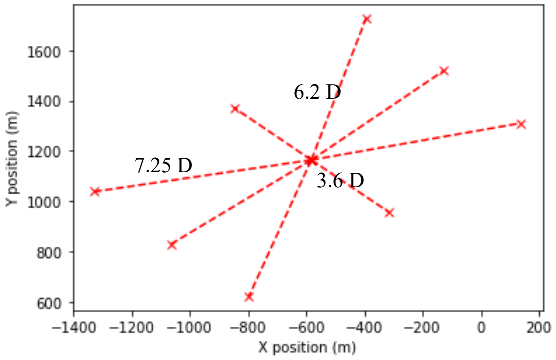
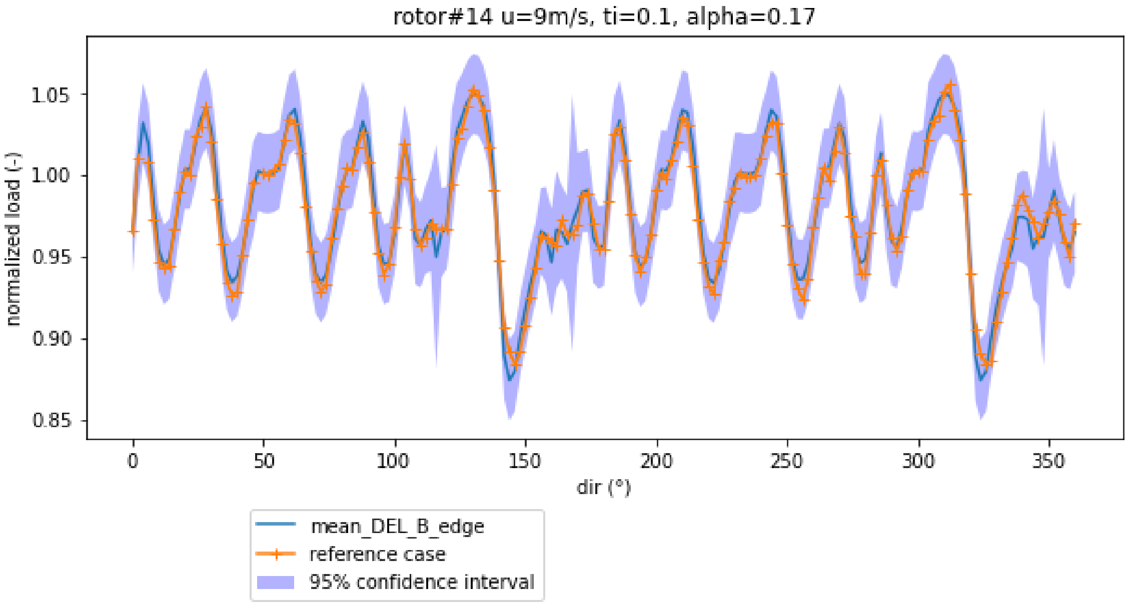
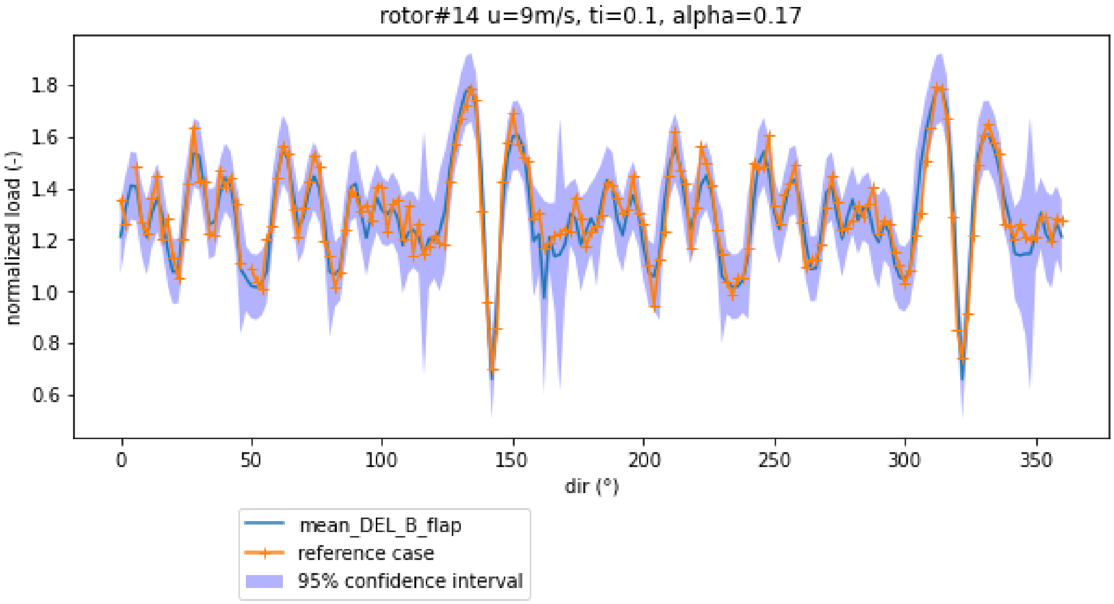
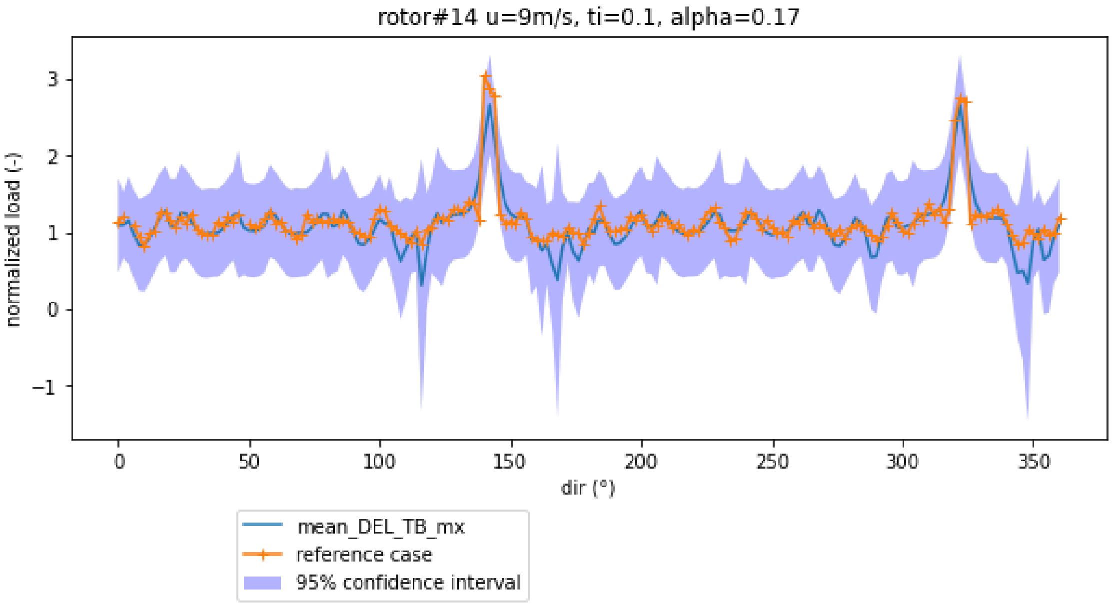


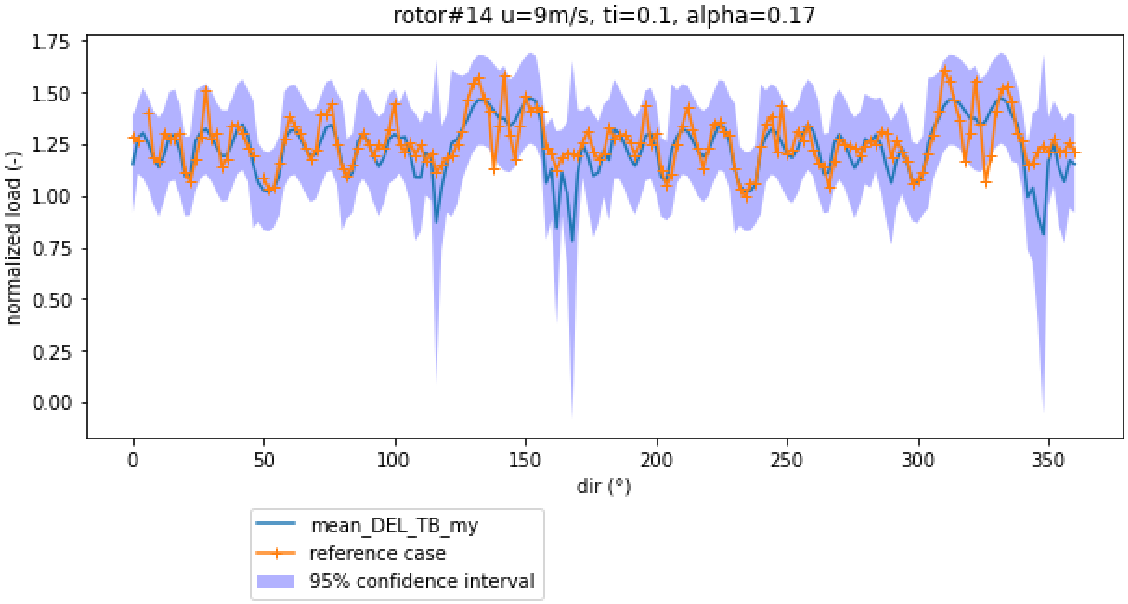
Figure 45 Turbine #14 and its 8 nearest neighbors

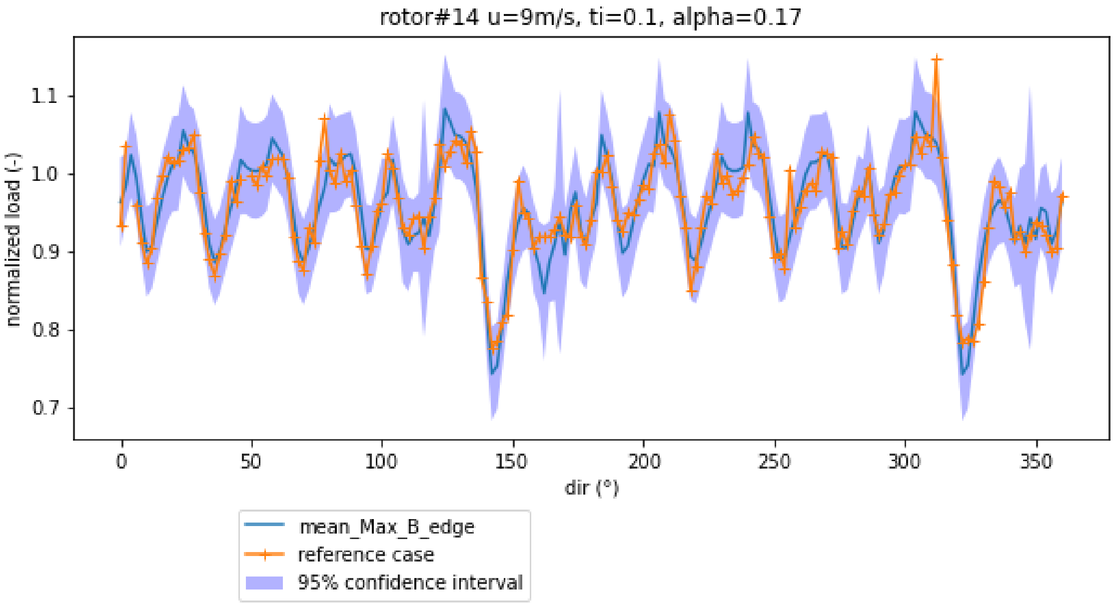
For that purpose, FarmShadow^{TM} simulations are executed for varying wind direction (from 0° to 360°, with a stepsize equal to 2°). For turbine #14 and for each direction, the mean velocity wind field profile is extracted and fitted with the model function to find the set of parameters. With this set of parameters as input, the surrogate models are used to predict fatigue loads and extreme loads. The loads are normalized with respect to their values in ambient conditions (no wake). The fatigue load predictions end the extreme load predictions are shown in Figure 46 and compared to the reference case, that is the output computed by the fully coupled simulator DeepLinesWind^{TM}.

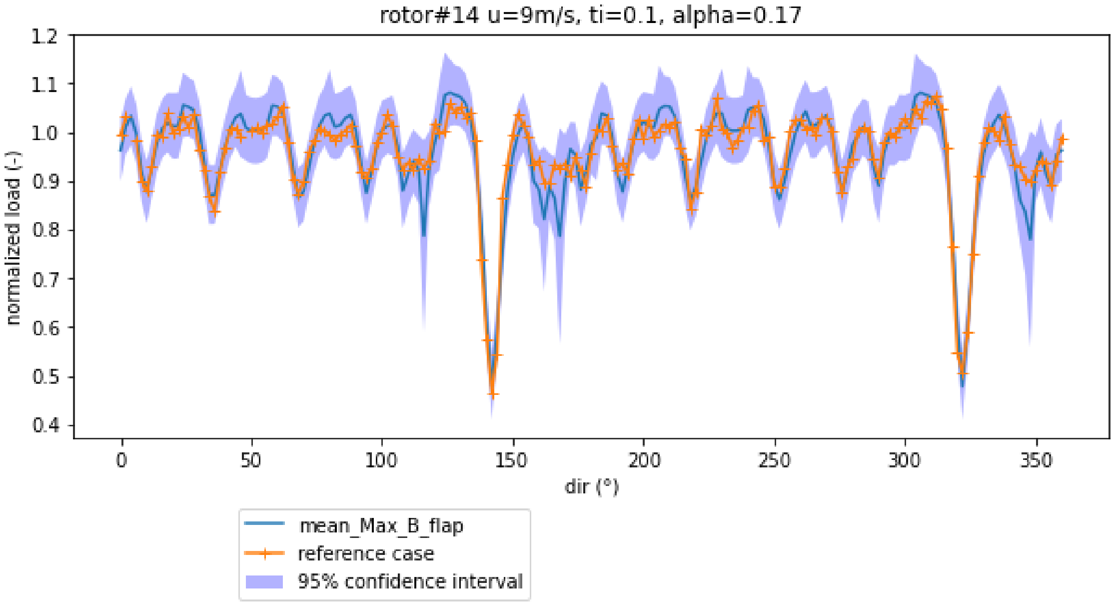





rotor#14 u=9m/s, ti=0.1, alpha=0.17
normalized load (-)
3
2
1
0
−1
0
50
100
150
200
250
300
350
dir (°)
mean_DEL_TB_mx
reference case
95% confidence interval


rotor#14 u=9m/s, ti=0.1, alpha=0.17
normalized load (-)
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
50
100
150
200
250
300
350
dir (°)
mean_DEL_TB_my
reference case
95% confidence interval


rotor#14 u=9m/s, ti=0.1, alpha=0.17
normalized load (-)
1.1
1.0
0.9
0.8
0.7
0
50
100
150
200
250
300
350
dir (°)
mean_Max_B_edge
reference case
95% confidence interval


rotor#14 u=9m/s, ti=0.1, alpha=0.17
normalized load (-)
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0
50
100
150
200
250
300
350
dir (°)
mean_Max_B_flap
reference case
95% confidence interval

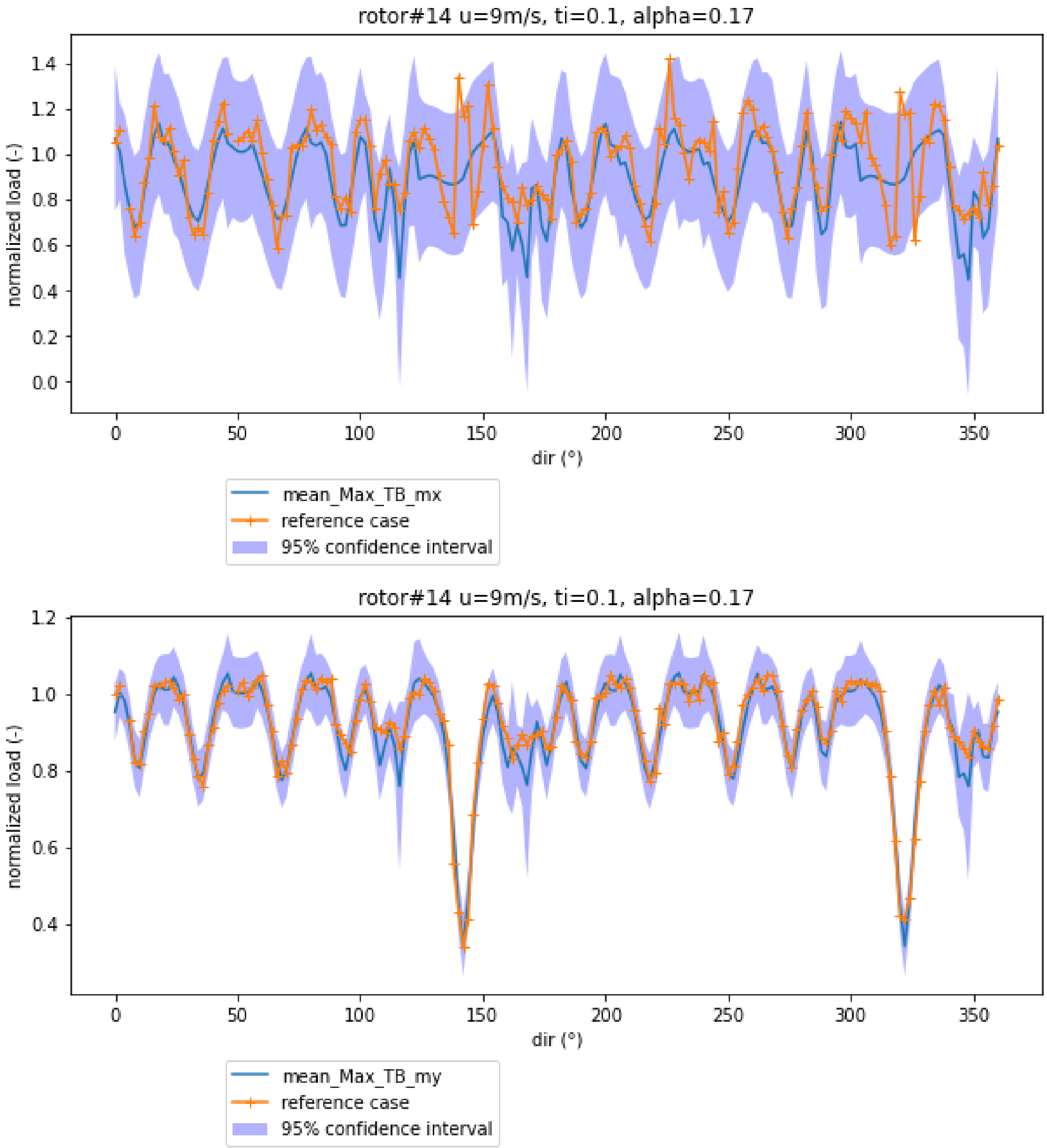


Figure 46 Fatigue load prediction and extreme load prediction vs reference case wrt. to wind direction

The DEL of the flapwise blade root bending moment exhibits M-shaped curve with two peaks surrounding a valley. The presence of the peaks, observed in other studies [19]–[21], may reflect the added turbulence due to the presence of the wake combined with the partial overlap of the rotor that creates an additional load variation on the blades. We can also notice an asymmetry in the shape: the left peak is a bit higher than the right peak. This is probably due to the rotational direction and the interaction of the wake with the shaft tilt: the load variation due to the shaft tilt may be slightly compensated by the presence of the wake depending on the relative position of the wake. More precisely, due to the tilt, when the blade goes upward, the blade faces away the wind direction and the apparent wind velocity is slightly lower: if the blade goes through the wake when rotating upward, the apparent wind velocity is much more lowered increasing the load fluctuations on the blade. Finally, the brutal decrease of the DEL when the wake is centered may be explained by the wake velocity deficit. Compared to the wake-free inflow, the partial wake conditions induce up to a 1.8 times increase of the DEL for a wind direction around 140°.

The M-shaped curve is also observed for DEL of the fore-aft tower bottom bending moment. It is not surprising as it is a load that is thrust dependent just as the flapwise blade root bending moment. The only difference is that the fore-aft tower bottom bending moment is probably much less sensitive to the wake overlap. Indeed, although individually, the blade load may exhibit a significant azimuthal variation, the fluctuations on three blades balance each other. The M shape is hence less pronounced for the fore-aft tower bottom bending moment. The same conclusions may be drawn for both the maximum flapwise and the fore-aft load except that the extreme loads seem to be more affected by the deterministic part of the load than by the stochastic part (in the extreme cases, the total Sobol indices related to the turbulence intensity are lower than in the fatigue cases).

The DEL of the side-side tower bottom bending moment exhibits a more complex behavior that may be the expression of the strong nonlinearity of this QOI wrt. the wind speed value. For the majority of QOI the relationship with the wind speed is monotonic (that is a monotically increasing function). This nonlinearity may be attributed to the drive train torsion resonance phenomenon for variable speed wind turbine (in addition, the aerodynamic damping is small for this mode). For the ambient conditions mentioned above (I_{u0}=0.1 and α=0.17), around 9 m/s, regardless the influence of wake induced added turbulence, the large velocity deficits may imply greater load fluctuations. (see Figure 47)

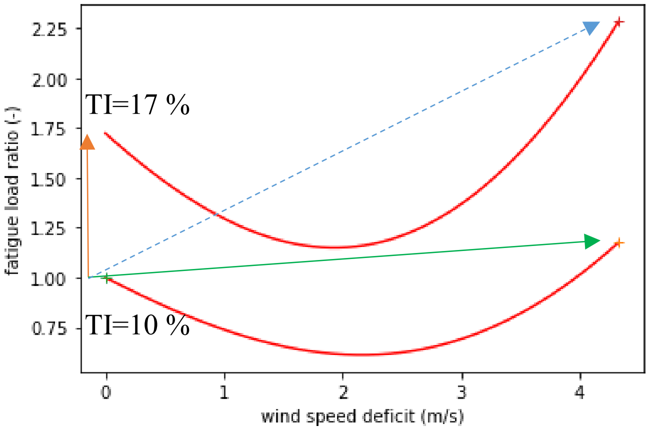


Figure 47 Side-side tower bottom bending moment fatigue load ratio wrt. ambient condition vs wind speed deficit (the wind speed deficit is related to u_0=9 m/s). The green arrow represents the increase of fatigue load due to the wind speed deficit. The orange arrow represents the increase of fatigue load due to added turbulence intensity

The DEL of the edgewise blade root bending moment is mainly dependent on the load variation due to the weight load and the rotational speed of the rotor: the dispersion wrt. to the wind direction of the edgewise blade root bending moment DEL is much smaller than the flapwise. The wake has a lower effect on this QOI. Nonetheless, it is worthy to note that variation of edgewise DEL when the wake hit the rotor may be explained by a compensation phenomenon: when the blade goes through the wake when rotating upward, the aerodynamic load that is acting in the opposite direction of the gravitational is smaller than in ambient condition and hence the mechanism of load compensation is lowered. On the other hand, if the blade goes through the wake when rotating downward, the aerodynamic load is acting in the same direction as the gravitational load and is lower than in ambient condition. Hence, the total load fluctuation is lowered. That is why the position of the edgewise blade root bending moment coincides with the position of the highest peak of the flapwise blade root bending moment. These observations are also valid for the edgewise extreme load.

4.7. Conclusion and perspectives

The objective of this task was to elaborate an efficient wake induced wind field parametrization and to provide a mapping between wake parameters and loads at specific wind turbines structural locations with the help of a surrogate model. The main difficulty is obviously to find a good compromise between wake representation accuracy and computational burden as the shape of the wake may be very complex inside a farm. The choice has been made here to work on parametrization based on a model function made of Gaussian shapes. This choice is clearly motivated by engineering analytical wake models. Considering the Teesside wind farm as a case study, the following conclusion may be drawn:

- ✓ the vast majority of the wake profiles for the Teesside wind farm study case may be quite well represented with a model function made at most of 2 Gaussians RBF functions. Out of a total of 26 973 wake profiles, almost 87 % may be fitted with the Gaussian based model function with a NRMSE lower than 10 %.
- ✓ the parametrization approach proposed in this section offers a farm independent parametrization that is based on the shape of the wake velocity deficit only. It was found that at most 6 parameters were required to fit the wake velocity field
- ✓ the association of this parametrization with a kriging based surrogate modeling has permitted to build a fatigue and extreme loads mapping exhibiting reasonable performance in terms of NRMSE. In particular, the surrogate model is able to reproduce some typical load variations when the wake interacts with the wind turbine. Besides, the global sensitivity analysis has shown the primary influence of free stream wind speed and turbulence intensity for both fatigue and extreme loads.

Some improvements of the computational process may be yet explored:

- ✓ since the addition of a third Gaussian is not expected to bring some fitting enhancement, the fitting process may be improved focusing on two axes:
    - o The suboptimality issue revealed by the accumulation at the bounds of the parameters range
    - o The enrichment of the model function with additional parameters (super Gaussian exponent, covariance matrix, z location of the center)
- ✓ regarding the surrogate modeling process, the size of the training data may be increased in order to lower the NRMSE values. Besides, it could be profitable, in case of larger training data, to test the performance of other surrogate technique, such as neural network.
- ✓ finally, concerning the parametrization approach, the impact of averaging turbulence intensity over the rotor should be evaluated by means of a DEL comparison on a specific case between DLW simulations with the actual wind field profile (including both turbulence intensity and velocity deficit spatial distribution) and the fitted one.

References in Section 4

[1] IEC, « IEC 61400-1:2019 - Wind energy generation systems - Part 1: Design requirements ». 2019.

[2] S. Tronæs Frandsen et Risø National Laboratory, « Turbulence and turbulence-generated structural loading in wind turbine clusters », Risø National Laboratory, Roskilde, 2007.

[3] F. Blondel et M. Cathelain, « An alternative form of the super-Gaussian wind turbine wake model », *Wind Energy Sci.*, vol. 5, nº 3, p. 1225-1236, sept. 2020, doi: 10.5194/wes-5-1225-2020.

[4] T. Göçmen *et al.*, « FarmConners Wind Farm Flow Control Benchmark: Blind Test Results », Wind and turbulence, preprint, févr. 2022. doi: 10.5194/wes-2022-5.

[5] M. Bastankhah et F. Porté-Agel, « A new analytical model for wind-turbine wakes », *Renew. Energy*, vol. 70, p. 116-123, oct. 2014, doi: 10.1016/j.renene.2014.01.002.

[6] PRINCIPIA, « DEEPLINES WIND™ ». https://www.principia-group.com/blog/product/deeplines-wind/

[7] T. Ishihara et G.-W. Qian, « A new Gaussian-based analytical wake model for wind turbines considering ambient turbulence intensities and thrust coefficient effects », *J. Wind Eng. Ind. Aerodyn.*, vol. 177, p. 275-292, juin 2018, doi: 10.1016/j.jweia.2018.04.010.

[8] N. Dimitrov, « Surrogate models for parameterized representation of wake-induced loads in wind farms », *Wind Energy*, vol. 22, nº 10, p. 1371-1389, oct. 2019, doi: 10.1002/we.2362.

[9] « MOLHD: Multiple Objective Latin Hypercube Design ». https://cran.r-project.org/web/packages/MOLHD/index.html

[10] M. D. Morris et T. J. Mitchell, « Exploratory designs for computational experiments », *J. Stat. Plan. Inference*, vol. 43, nº 3, p. 381-402, févr. 1995, doi: 10.1016/0378-3758(94)00035-T.

[11] T. Juul Larsen, A. Melchior Hansen, Ris?? National Lab, et W. E. D. DTU, *How 2 HAWC2, the user's manual.* 2007.

[12] B. J. Jonkman, « Turbsim User's Guide: Version 1.50 », NREL/TP-500-46198, 965520, sept. 2009. doi: 10.2172/965520.

[13] M. Matsuichi et T. Endo, « Fatigue of metals subjected to varying stress », 1968.

[14] C. E. Rasmussen et C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.

[15] V. Dubourg, « Adaptive surrogate models for reliability analysis and reliability-based design optimization », p. 309.

[16] A. Saltelli, S. Tarantola, et K. P.-S. Chan, « A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output », *Technometrics*, vol. 41, nº 1, p. 39-56, févr. 1999, doi: 10.1080/00401706.1999.10485594.

[17] M. Baudin, A. Dutfoy, B. Iooss, et A.-L. Popelin, « OpenTURNS: An Industrial Software for Uncertainty Quantification in Simulation », in *Handbook of Uncertainty Quantification*, R. Ghanem, D. Higdon, et H. Owhadi, Éd. Cham: Springer International Publishing, 2017, p. 2001-2038. doi: 10.1007/978-3-319-12385-1_64.

[18] T. Homma et A. Saltelli, « Importance measures in global sensitivity analysis of nonlinear models », *Reliab. Eng. Syst. Saf.*, vol. 52, nº 1, p. 1-17, avr. 1996, doi: 10.1016/0951-8320(96)00002-6.

[19] I. Reinwardt, L. Schilling, D. Steudel, N. Dimitrov, P. Dalhoff, et M. Breuer, « Validation of the dynamic wake meandering model with respect to loads and power production », *Wind Energy Sci. Discuss.*, vol. 2020, p. 1-27, 2020, doi: 10.5194/wes-2020-126.

[20] I. Reinwardt, N. Gerke, P. Dalhoff, D. Steudel, et W. Moser, « Validation of wind turbine wake models with focus on the dynamic wake meandering model », *J. Phys. Conf. Ser.*, vol. 1037, p. 072028, juin 2018, doi: 10.1088/1742-6596/1037/7/072028.

[21] J. M. Rinker, E. Soto Sagredo, et L. Bergami, « The Importance of Wake Meandering on Wind Turbine Fatigue Loads in Wake », *Energies*, vol. 14, nº 21, 2021, doi: 10.3390/en14217313.

5. Comparison between wake parameterization method outputs

The surrogate modelling approaches presented in Section 3 and Section 4 provide two different workflows towards completing the same task – prediction of loads and power outputs on an individual turbine in an arbitrary wind farm, given the ambient conditions and the wind farm layout as inputs. Apart from the individual model performance assessment done earlier, additional insights can be gained by comparing the two approaches on the same case, so that we can see eventual patterns and similarities. We made such a comparison for a single case with 9m/s ambient wind speed which was chosen because it is within the region with maximum thrust of the turbines, and leads to strong wake deficits that are more challenging for the models to predict as seen in Sections 3 and 4. The turbine of interest is again T14 at the Teesside wind farm, its location indicated in Figure 3. Other environmental conditions considered include turbulence intensity of 10% and wind shear exponent of 0.17. Load simulations of T14 with Hawc2 and the Dynamic Wake Meandering model with the entire wind farm included as wake sources are carried out for wind directions from 0° to 358° in steps of 2 degrees. In order to capture the realization-to-realization uncertainty and obtain smoother load polars, 12 random realizations are simulated at each wind direction, and the statistics (mean and average over all realizations) of the quantities of interest are computed. Figure 48 and Figure 49 show the comparison between the surrogate model predictions and the reference simulations, for two illustrative load channels – the blade root flapwise and edgewise bending moment DELs respectively. In these comparisons, "DWM surrogate" stands for the model predictions based on the procedure developed by DTU and described in Section 3 of this report, while "DL surrogate" stands for "DeepLines Wind surrogate" - the model predictions based on the procedure developed by IFPEN and described in Section 4 of the report. We can see that for the big wake peaks at about 120° and 310° the DL surrogate predicts somewhat smaller load variation than the DWM surrogate: the load increase due to half-wake situations is smaller, and the load decrease in the center due to the reduced equivalent wind speed in full-wake situations is bigger. This is likely a feature of the wake model and not the surrogate itself. At the same time, for less strong wake conditions (as e.g. in the range of 0° to 100°, with fewer disturbing turbines that are farther away), the DL surrogate seems to give more consistent results than the DWM surrogate. These differences may be explained by two factors: 1) the use of Kriging in the DL surrogate where the data points known from the model training are utilized in the predictions and the error close to the known points is reduced; and 2) the DWM approach is subject to additional uncertainties due to the randomness in the turbulence realizations and the wake meandering – meaning that more data may be required to further improve the results.

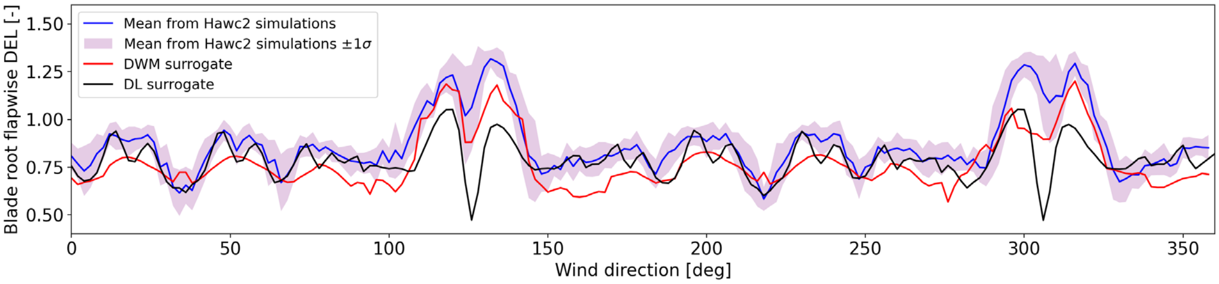


Figure 48 Comparison of surrogate model predictions of blade flapwise bending DEL against Hawc2 aeroelastic simulations.

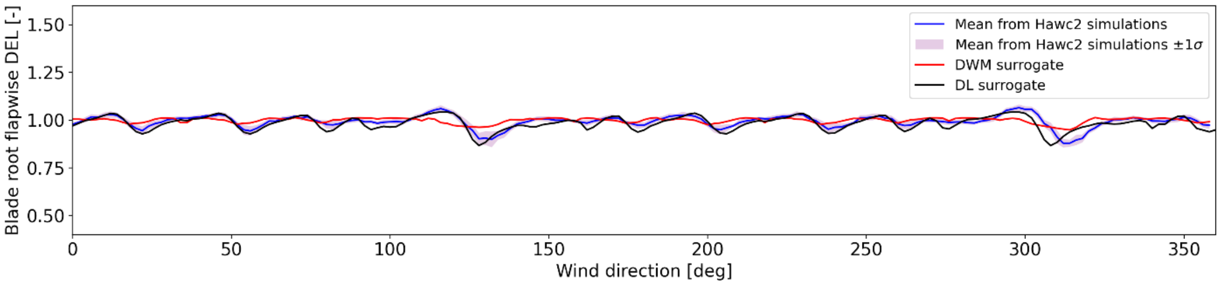


Figure 49 Comparison of surrogate model predictions of blade edgewise bending DEL against Hawc2 aeroelastic simulations.

Based on the above comparison, and on the individual results presented in Section 3 and 4, we can draw the following conclusions common for the two surrogate modelling workflows:

- Overall, we get consistent picture on the sensitivity with both approaches – that wind shear has only minimal influence on the Sobol indices, and that wind speed has the greatest influence
- The performance of both models is better in predicting the mean of extremes rather than DELs
- The approach based on describing the wake deficit with shape functions ("DL surrogate") gives smoother shapes – which could be a combination of the fact that it is using Kriging, and that using a stationary wake profile probably reduces some of the uncertainty (which is present with the DWM-based approach due to the wake meandering).

6. Transient wind event simulations using a constrained turbulence approach

6.1. The Hipersim/turbgen tool for turbulence box generation and constrained turbulence simulations

The turbulence box generation tool used in the Hiperwind project is called Hipersim, and is available as a free Python package (pypi.org/hipersim). It is specifically the Turbgen module in Hipersim that provides the turbulence-related functions.

The basic functionality of the Turbgen module is to generate wind fields following the Mann (1994) [1] spectral tensor definition, based on the algorithm outlined in Mann (1998) [2]. Another major function is to generate turbulence boxes with embedded constraints that specify the exact values of the field at a list of points; for example, such constraints can represent measured time series that we want to recreate in the otherwise random turbulence box. The constrained simulation algorithm is described in more details in Dimitrov & Natarajan [3].

Installing Turbgen is done by installing the parent Hipersim package:

```
pip install hipersim
```

Importing Turbgen into the Python environment:

```
from hipersim.turbgen import turbgen
```

The wind field generation function takes the following basic inputs:

- N_x: Dimension of the turbulence box in longitudinal (x) direction (default = 8192)
- N_y: Dimension of the turbulence box in transverse horizontal (y) direction (default = 32)
- N_z: Dimension of the turbulence box in transverse vertical (z) direction (default = 32)
- dx: Spacing in meters between data points along the x-coordinate (default = 1m)
- dy: Spacing in meters between data points along the y-coordinate (default = 1m)
- dz: Spacing in meters between data points along the z-coordinate (default = 1m)
- L: Mann model turbulence length scale parameter L (Mann, 1994) (default = 29.4)
- Γ : Mann model turbulence anisotropy parameter Γ (Mann, 1994) (default = 3.9)
- $alphaepsilon$: Mann model turbulence parameter $\alpha\varepsilon^{2/3}$ (Mann, 1994) (default = 1.0)
- $HighFreqComp$: Defines whether high-frequency compensation is applied to the turbulence generation as defined in (Mann, 1998). Default = 0 (no high-frequency compensation).

Additional details about the Hipersim tool can be found in its use manual [4].

Below, a few figures illustrate the results produced by the Turbgen tool, and provide basic means of validation. Figure 50 shows an example turbulence field generated by the tool with default Mann model parameters. The spectra and coherence at different spatial separations of generated wind fields are compared to the theoretical coherences and spectra in Figure 51. Finally, an illustration of how constraints are applied and what are the changes introduced to the resulting field, is shown in Figure 52 (see figure caption for more details about the plot contents).

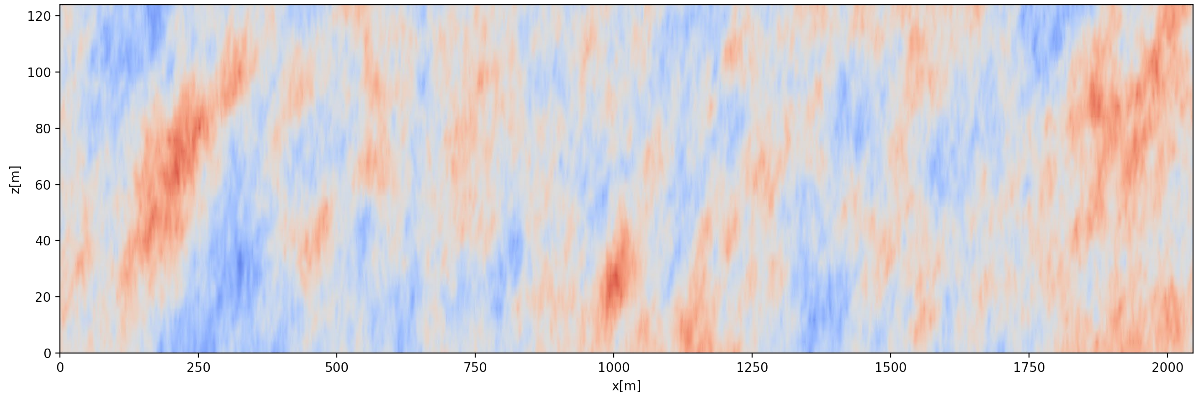


Figure 50 Example turbulence field generated by Turbgen, based on the Mann turbulence spectrum with $L = 30$, $\Gamma = 3.9$ *and* $\alpha\varepsilon^{2/3} = 1$

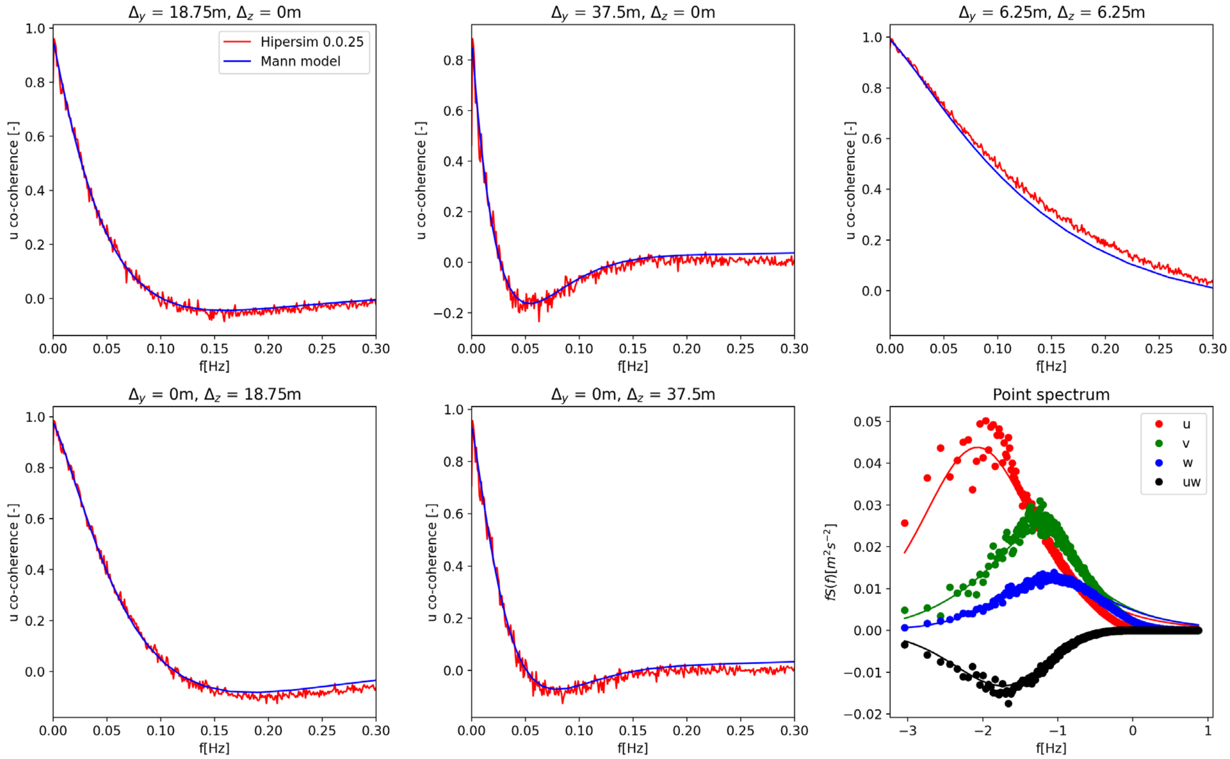


Figure 51 Coherence and spectrum of turbulence boxes generated with the Hipersim/Turbgen, compared to theoretical coherences and spectra. The curves denoted as "Hipersim" are ensemble averages from 10 turbulence boxes of size 8192x32x32.

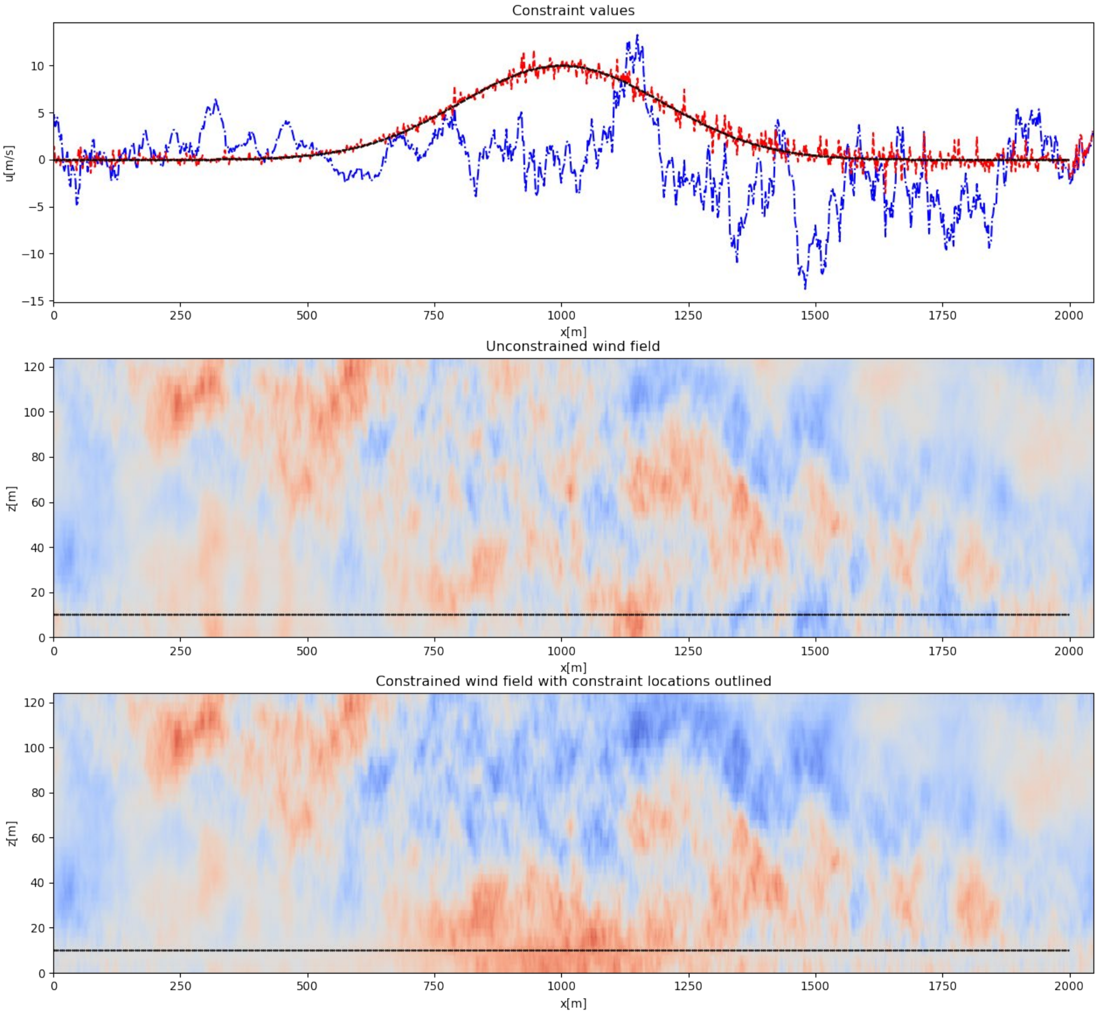


Figure 52 Illustration of the effect of applying constraints to a random wind field with size 2048m x 128m x128m. Top plot: blue curve shows the time series of the random turbulence field before the constraints are applied at a location of y = 60m and z = 10m. Black curve in the top plot shows the required constraints values, and red curve shows the time series after the constraints have been applied. Middle plot: a cross-section of the unconstrained field in the x-z plane at y = 60m. The dashed line indicates the location of the constraints. Bottom plot: a x-z cross-section of the field after the constraints have been applied. The area with increased wind speed that satisfies the constraint requirements is visible with stronger red shading.

6.2. Impact of transient event profiles on turbine loads

Introduction

The Hipersim turbgen tools allows us to recreate the flow structures seen in measurements within a synthetic turbulence box. There after we can use this turbulence box to calculate the loads produced from these events. The primary focus of this section is to assess how well the constrained turbulence boxes recreated the events seen in measurements and to compare the loads from these events to the loads from the synthetic events required in the IEC 61400 standards.

This section starts by describing the overall methodology and workflow. Second, we compare the synthetic turbulence with the meteorological data. Finally we compare the different loads.

Methodology

The constrained turbulence is based on 304 unique 10 minute time series of u and v wind measurements at the Høvsøre test station in Denmark, at nominal wind speeds 8.5, 9.5, 10.5, 11.5, 12.5, 13.5, 14.5, 15.5 and 16.5 in m/s. These time series were selected as the top 10 for axial acceleration (i.e. du/dt), direction change (dφ/dt) at 100m and 160m, for each wind speed. In some cases, a single time series could be in the top 10 in multiple categories. Further note that the time series were selected based on filtered accelerations and direction changes. The filtering is based on the second order low pass Butterworth filter, with a cut-off frequency of 0.1Hz. The constrained turbulence is based on the unfiltered results, down-sampled to 1Hz.

In addition to the time series, the length scale L, and the $\alpha\varepsilon^{2/3}$ and Γ parameters for the Mann turbulence model were also estimated based on the time-series. For each time-series, two sets of parameters were provided. The first is based on the data unfiltered. Since this included the significant events, the turbulent length scale was estimated to be very large and unrealistic. The second set of parameters was based on time series that were high-pass filtered to remove spatial fluctuations on mesoscale level (2km or larger), and as a result better reflect the background turbulence. This high-pass filtering procedure is described in more details in [8]. The two sets of parameters are referred, in accordance to the length scale, as the large and small parameters respectively.

The first stage in the investigation was generating the turbulence boxes. The constraints were applied in 4 different configurations. Figure 53 shows the different constraint locations in relation to the rotor disk. The first configuration used the small parameters to define the background turbulence. Then the constraints were applied using the large parameters. The constraints were applied at the center of the turbulence box at the respective heights of the measurements (100m 160m). These locations correspond to the black circles in Figure 53. The second configuration is identical to the first, except the small Mann parameters were used for both the background turbulence and the application of the constraints. The measurements do not give any indication to the scale of the structures, so the third configuration considered the situation where the event would apply to the whole rotor simultaneously. The same constraints were applied at two lateral locations, offset from the center, with the large Mann parameters. The locations correspond to the red and blue locations in Figure 53. Finally, a fourth configuration considers the effect on loads of the blades periodically slicing through a turbulence structure to the side of the rotor disk. In this configuration, the constraints are applied laterally offset to one side, with the small Mann parameters, so they manifest as smaller local structures. This corresponds to the red circles only.

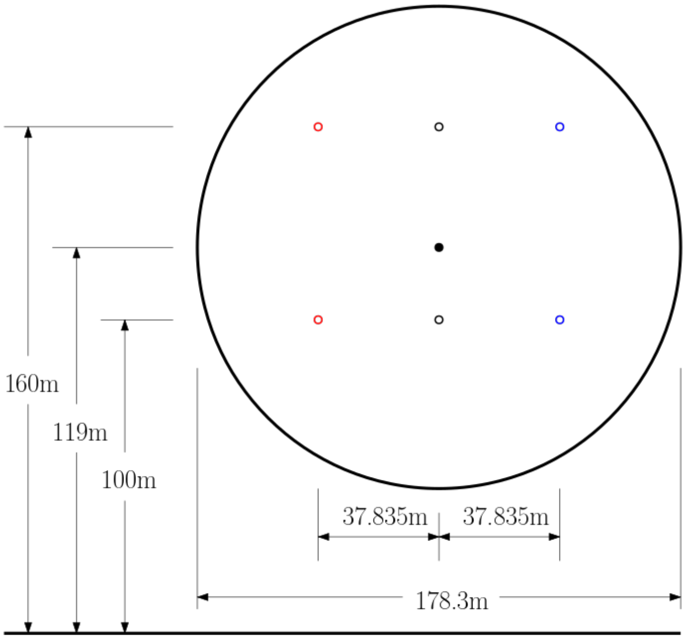


Figure 53 The different locations for turbulence constraints in relation to the rotor disk.

The data provided measurements from two heights and thus give an indication of the shear. The effect of shear in the constrained simulations was simulated in two different ways. The first assumed that the application of the constraints would correctly introduce shear into the turbulent boxes, thus the HAWC2 simulations would not need to add shear to the flow. A second treatment to shear subtracted the shear from the time series before applying the constraints. Then included the same shear to the HAWC2 simulations. This second way is useful to ensure shear is applied throughout the whole domain and not just near the constraint locations.

In summary, the different considerations for DOE of constrained turbulence loads simulations are as follows:

- Different configurations for applying the constraints:
  - Rotor centered and constrained with large Mann parameters
  - Rotor centered and constrained with small Mann parameters
  - Offset laterally on both sides and constrained with large Mann parameters. Meant to represent large coherent events
  - Offset laterally on one side and constrained with small Mann parameters. Meant to simulate the effect on loads of a blade slicing through a significant turbulent structure
- Different configurations for accounting shear
  - The effect of shear is introduced through the constraints. The constraint values are not offset according to a prescribed shear and the HAWC2 simulations do not add the prescribed shear to the load simulations.

- The effect of shear is introduced through the load simulations. The prescribed shear is subtracted from the constraint values when the constraints are applied and then reintroduced in the HAWC2 loads simulations.

The DTU 10MW Reference turbine was chosen to assess the impact of these constrained simulations. The rotor has a diameter of 178.3m and a hub height of 119m, so the both the measurement heights intersect with the rotor disc. Furthermore, this model was chosen because the controller (including yaw control) is open source and available. More details on this turbine can be found here [5].

Load calculations based on the IEC 61400 standard [6] are used as a reference to compare the loads from the constrained turbulence. The EOG, EDC, ECD extreme events described in the standard are similar to the significant wind events in the time series according to how the wind conditions are changing. The IEC 61400 design standards require that these events are simulated with a shutdown and power-loss simultaneously. However, these events were simulated without the shutdown or power-loss so that effect of these extreme events can be compared to the significant wind events directly.

An important parameter in the IEC load simulations is σ_1, it is related to the turbulence intensity for the wind. It also defines the magnitude of the EDC and EOG events. This value can also be determined from site data directly by taking the standard deviation of a wind measurement. These values where then grouped by height and wind speed and averaged. Then these averages were interpolated to the hub-height to define a site-specific σ_1 value for each wind speed.

The σ_1 can also be estimated from the turbulence length scale L and $\alpha\varepsilon^{2/3}$,, derived from the filtered time series. This defines a second set of σ_1 values. The same procedure of averaging and interpolating was applied to this data, to define a second site-specific σ_1 value for each wind speed. Those values can be seen Figure 54, compared to the recommended values for different IEC 61400 sites.

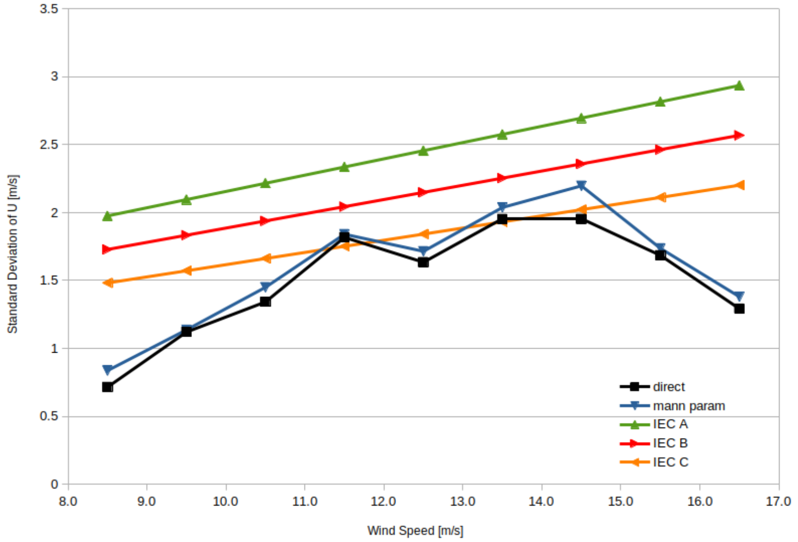


Figure 54 Site-specific turbulence vs. IEC reference classes

The IEC wind conditions themselves are based on several parameters. Different variations of the IEC load cases were simulated to investigate the different effects of these parameters and to explore how the different site specific information effected the loads. The parameters of interest are as follows:

- **Events vs. Normal Operation:** The events only occur for a short time frame within the 10 minute simulations. To understand how momentary events effect the load statistics, a series of normal operation load cases was simulated.
- **Shear vs. No Shear:** The default condition used a power law shear profile with an exponent of 0.2. A second set of results simulated the wind without shear.
- **Turbulence Intensity:** The effect of turbulence is explored by varying the turbulence intensity. There are a set of simulations without turbulence, then with the IEC NTM and ETM. To explore the effect of site-specific turbulence intensity, two additional sets of IEC load simulations are calculated using the two site-specific σ_1 values calculated above.
- **Site Specific Mann Parameters:** A pair of IEC load calculations were executed to compare the effect of the parameters to generate the turbulence boxes (i.e. turbulent length scale, alpha_epsilon and gamma). The first uses the recommended values from the IEC 61400 standards and second uses the estimated values from the time-series, using the background values.
- **Magnitude of the EOG and EDC events**: The σ_1 values also defines the magnitude of the EOG and EDC. Another set of IEC load calculations was executed to compare the IEC recommended values with the values derived from the σ_1 values calculated above.

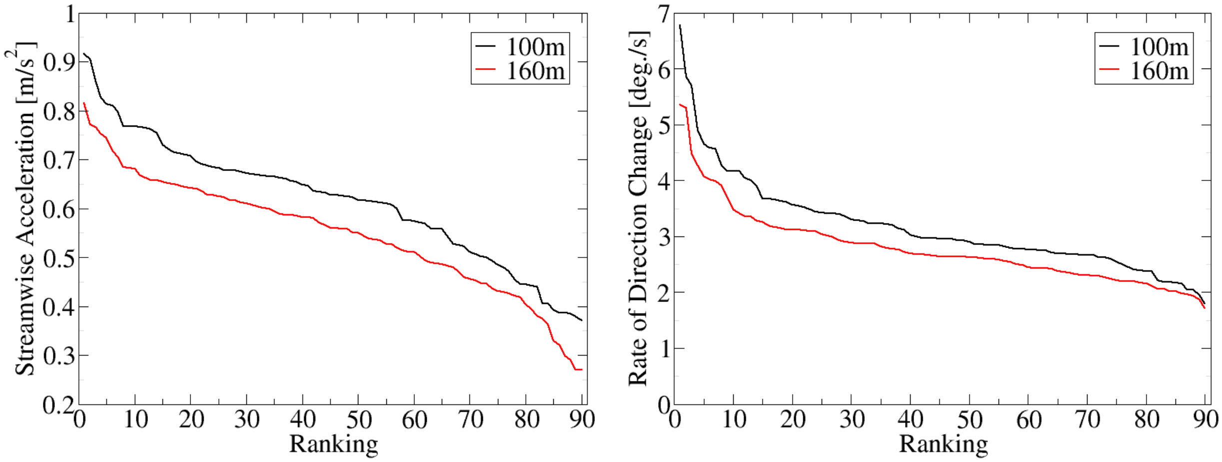
To explore the effects of the different significant wind events, the loads were calculated with each constrained turbulence box. To ensure that the events are reproduced within the load simulations, the turbulence boxes are not scaled to achieve a specified turbulence intensity. Thus, the turbulence intensity in the simulation it based on the results of the turbulent boxes themselves.

The results of these significant wind events are grouped according to the type of event. Significant changes in velocity at 100m and 160m are compared to EOG extreme events. Similarly, significant changes in direction at 100m and 160m are compared to EDC. Finally, time series that had a significant change of speed along with a significant change in direction, both at either height, were grouped and compared to ECD extreme events.

For both the IEC load calculations and significant wind load calculations, 10 minute simulations were run using HAWC2 [7].

Constrained Turbulence Boxes

In order to assess the loads from significant wind events, it is important to assess how well the constrained turbulence tools can reproduce these significant wind events. As stated earlier, the significant wind events were selected based on low-pass filtered stream-wise and directional accelerations. Figure 55 below shows P99% of filtered accelerations for time series used in this study. See the Hiperwind D2.3 report [8] for more details.



99% Quantile of Low-pass Filtered Stream-wise Acceleration for the times series

99% Quantile of Low-pass Filtered Rates of Direction Change for the time series

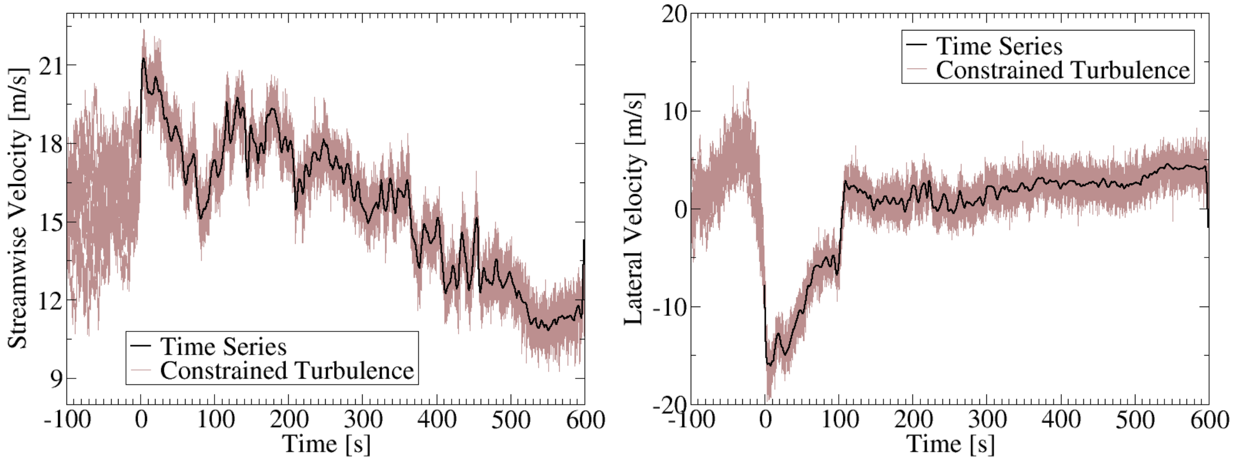
Figure 55 Low-pass filtered acceleration distributions.

To validate the constrained turbulence, two time series were selected for closer study. The statistics and properties are given in the following table. The 2015 event was selected since it had a strong stream-wise acceleration, while the 2014 event was selected due to the high rate of direction change. For the remainder of this section, the two events will be referred to as the acceleration and direction event respectively.

Table 15 Overview of two example time series classified as a stream-wise acceleration event and a direction change event.

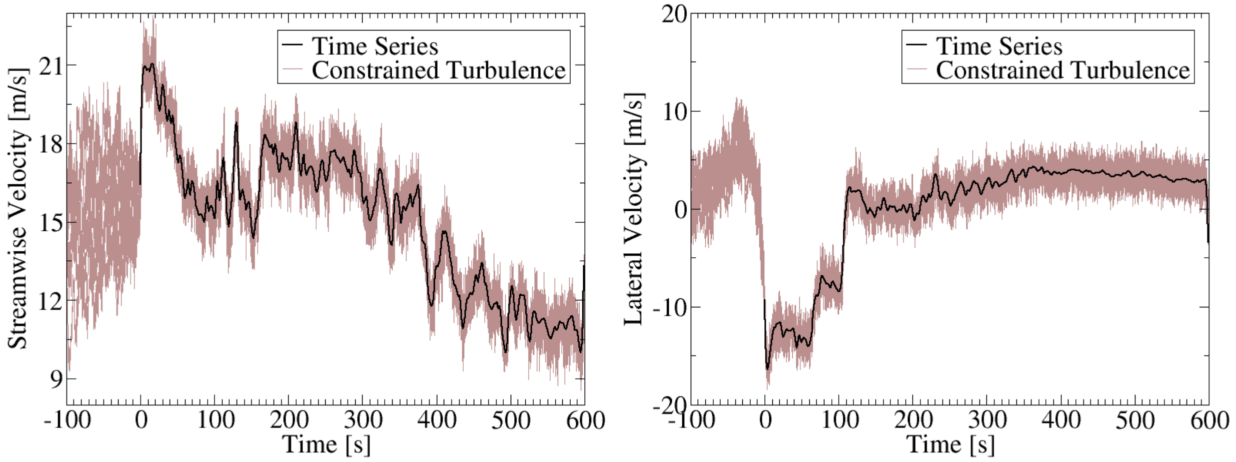
Time of Time Series	Dec. 11 2015, 8:50	Nov. 2 2014, 22:30
Nominal Wind Speed	15.5 m/s	12.5 m/s
Event of Interest	Acceleration	Direction
99% Quantile of stream-wise acceleration at 100m	0.903 m/s2	0.825 m/s2
99% Quantile of stream-wise acceleration at 160m	0.811 m/s2	0.628 m/s2
99% Quantile of rate of direction change at 100m	2.51 deg./s	6.77 deg./s
99% Quantile of rate of direction change at 160m	2.69 deg./s	3.36 deg./s
Time of significant stream-wise acceleration at 100m	597s	50s
Time of significant stream-wise acceleration at 160m	597s	200s
Time of significant of rate of direction change at 100m	1s	106s
Time of significant of rate of direction change at 160m	597s	106s

To start the assessment, the plots in Figure 56 show how the constrained turbulence was able to reproduce the wind velocity as the time series. These plots show the full time horizon of the simulation, where the first 100s is unconstrained and provides the simulation time to initialize the loads simulation. The constrained turbulence starts at t=0s. Ten different realizations of the constrained turbulence were generated. All the constrained turbulence data is shown at once in the brown curves.



Comparison of the stream-wise velocity at 160m for the acceleration time series.

Comparison of the lateral velocity at 160m for the direction time series.



Comparison of the stream-wise velocity at 100m for the acceleration time series.

Comparison of the lateral velocity at 100m for the direction time series.

Figure 56 Example measured time series and their reproduction in turbulence boxes using the constrained turbulence approach.

The data within the synthetic turbulence box was processed to produce the filtered acceleration and direction change. Comparing this filtered data with the corresponding data with the time series will show how well the constrained turbulence boxes reproduce the filtered accelerations and rates of direction change. Table 16 shows the filtered acceleration for the boxes based on the acceleration event time series at the same time that the corresponding significant event was detected.

Table 16 Filtered acceleration statistics for multiple realizations of constrained turbulence boxes based on the same target time series. Acceleration is in m/s^2.

	100m	160m
Corresponding Time Series Data	0.907	0.81
Synthetic Data Seed 3470	0.525	0.511
Synthetic Data Seed 3471	0.435	0.533

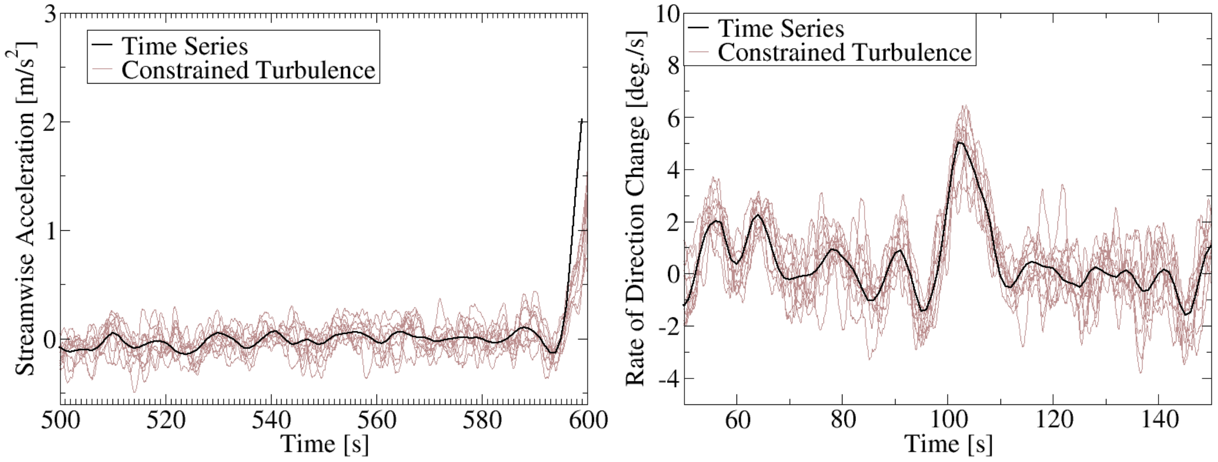
Synthetic Data Seed 3472	0.476	0.562
Synthetic Data Seed 3473	0.786	0.349
Synthetic Data Seed 3474	0.769	0.499
Synthetic Data Seed 3475	0.542	0.280
Synthetic Data Seed 3476	0.514	0.341
Synthetic Data Seed 3477	0.638	0.364
Synthetic Data Seed 3478	0.257	0.571
Synthetic Data Seed 3479	0.524	0.470
Minimum of Synthetic Data	0.257	0.280
Maximum of Synthetic Data	0.786	0.571
Average of Synthetic Data	0.547	0.448
Standard Deviation of Synthetic Data	0.156	0.105

Similarly, Table 17 shows the filtered rate of direction change within the synthetic data, at the same time of the significant event in the time series data.

Table 17 Statistics of filtered rates of direction change for multiple realizations of constrained turbulence boxes based on the same target time series. Rate units are in deg/s.

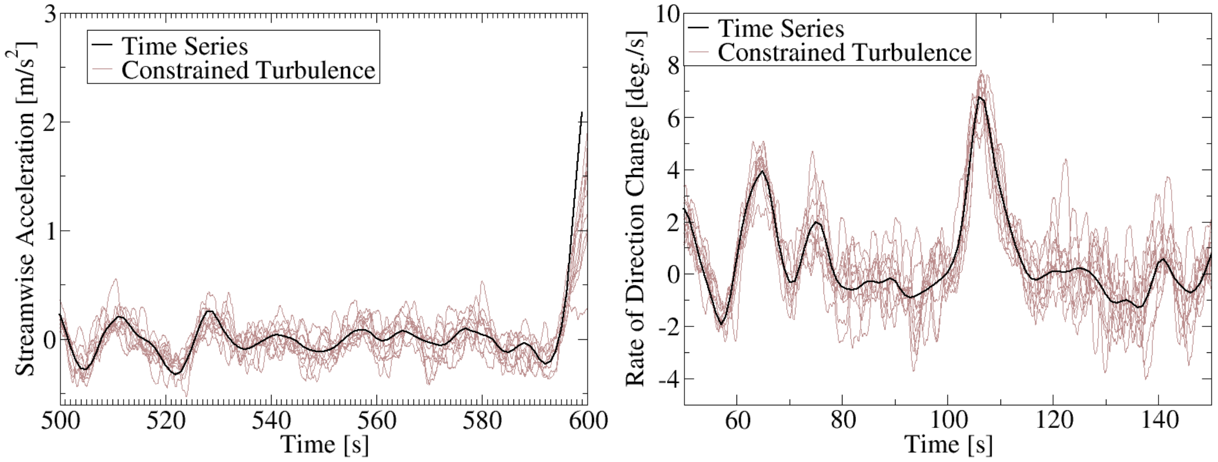
	100m	160m
Time Series Data	6.770	3.36
Synthetic Data Seed 2530	4.81	3.91
Synthetic Data Seed 2531	6.52	3.22
Synthetic Data Seed 2532	7.44	4.93
Synthetic Data Seed 2533	7.60	3.23
Synthetic Data Seed 2534	5.77	3.78
Synthetic Data Seed 2535	7.49	2.00
Synthetic Data Seed 2536	7.32	2.35
Synthetic Data Seed 2537	7.13	2.45
Synthetic Data Seed 2538	5.18	2.85
Synthetic Data Seed 2539	7.40	3.73
Minimum of Synthetic Data	4.81	2.00
Maximum of Synthetic Data	7.60	4.93
Average of Synthetic Data	6.67	3.24
Standard Deviation of Synthetic Data	1.04	0.88

The numerical data shows that the constrained turbulence data could not perfectly reproduce the same filtered accelerations or rate of direction changes. To further understand these discrepancies, the following plots in Figure 57 compare how filtered acceleration and rate of direction change vary over time.



Comparison between time series and synthetic turbulence of low-pass filtered stream-wise acceleration at 160m

Comparison between time series and synthetic turbulence of low-pass filtered rate of change in direction at 160m



Comparison between time series and synthetic turbulence of low-pass filtered stream-wise acceleration at 100m

Comparison between time series and synthetic turbulence of low-pass filtered rate of change in direction at 100m

Figure 57 Change of filtered stream-wise acceleration and rate of direction change: comparison between measured time series and synthetic turbulence realizations.

The time series show that the filtered gradient of the synthetic turbulence shows good agreement with the time series in the interior of the time series, but not at the boundary. The way that the gradient is calculated is based on taking the Fourier series, applying the filter and taking the gradient and then taking the inverse Fourier transform. An implied assumption of a Fourier series is that the data is periodic. Since the data is not periodic, this assumption is a likely source of these errors. Hence, we can conclude that constrained turbulence reproduces the acceleration properties in the measurements.

Another important property is the scale and coherence of the constrained flows. The following plots show contour plots of the stream-wise velocity on the x-y plane at an altitude of 100m for the large Mann parameter configurations. The velocity plot in Figure 58 clearly shows the effect of the constraint with a large structure starting at 1800m. The plot is followed by the corresponding filtered acceleration plot in Figure 59. The acceleration plot clearly shows how the constraint generates thin region of high acceleration. Another feature is that the effect of the constraint on velocity diffuses out over the whole rotor, while the acceleration gets attenuated over shorter length scales.

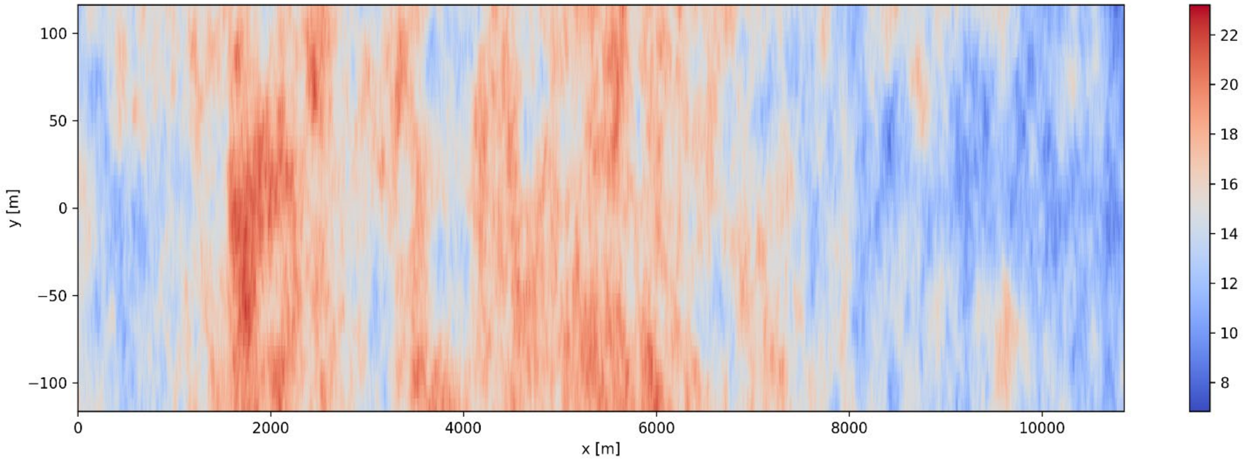


Figure 58 Stream-wise velocity [*m/s*] *at 100m for the large Mann parameter configuration*

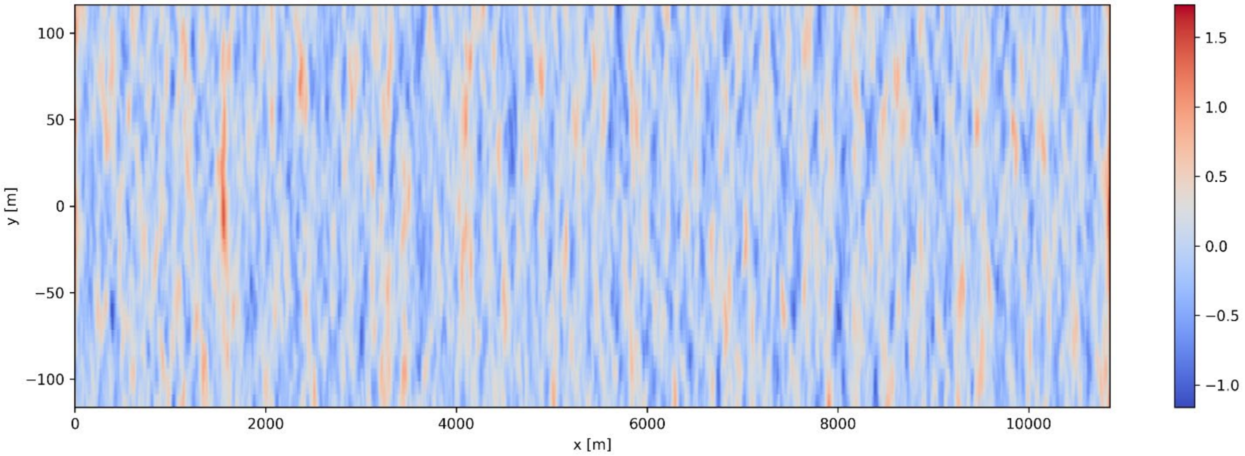


Figure 59 Filtered stream-wise acceleration $[m/s^2]$ *at 100m for the large Mann parameter configuration*

The following plot shows the same stream-wise velocity plot for the constrained turbulence box based on applying the constraints with the small Mann parameter configuration. Comparing with the large Mann parameter result shows that there is little difference to the velocity solution. The corresponding stream-wise acceleration plot is given below. Comparing this plot with the large background version shows that applying the constraints with large Mann parameters lead to stronger accelerations in the flow.

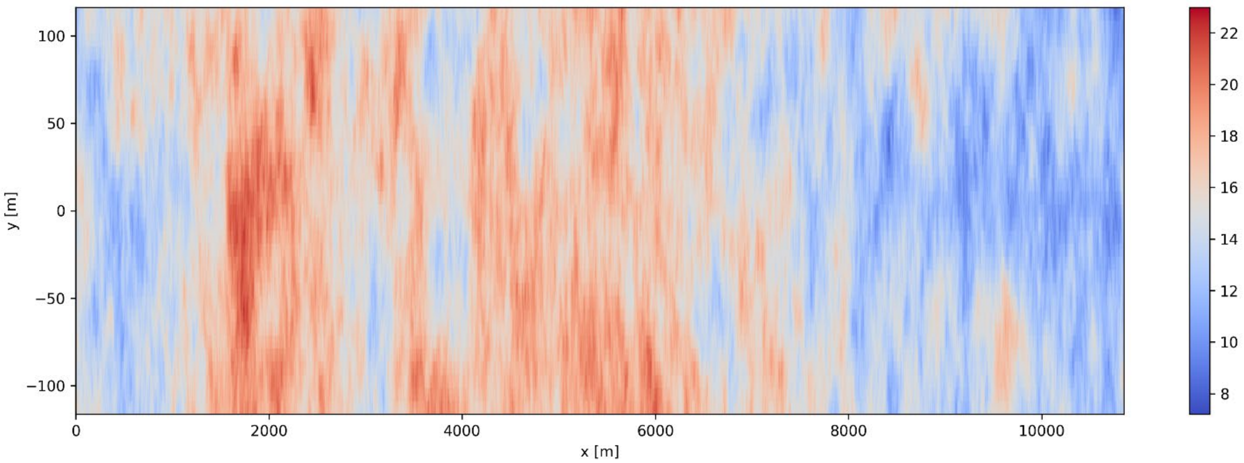


Figure 60 Stream-wise velocity [*m/s*] *at 100m for the small Mann parameter configuration*

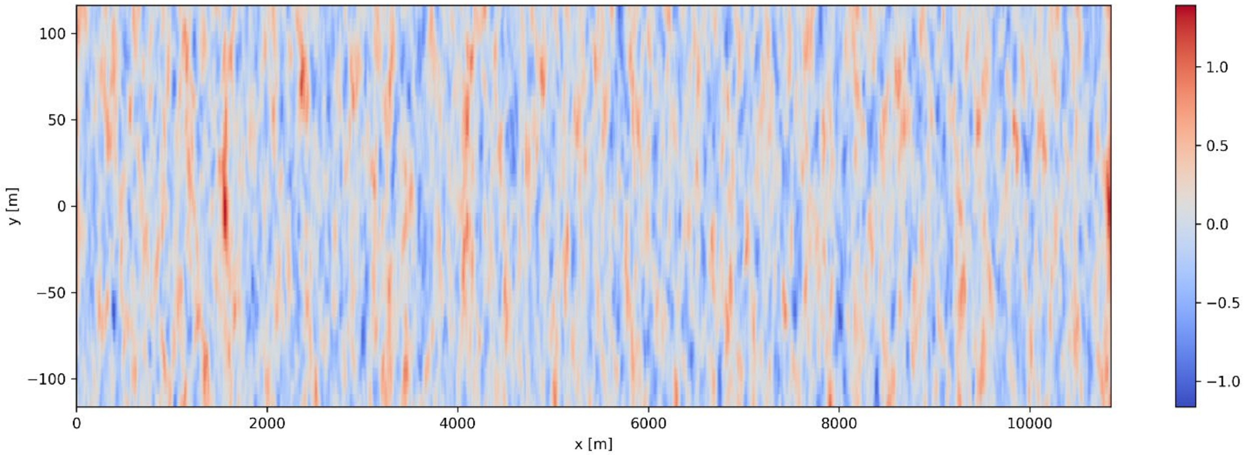


Figure 61 Filtered stream-wise acceleration [m/s^2] *at 100m for the small Mann parameter configuration*

The coherent constraint configuration is another technique for controlling the scale of turbulent structures. The stream-wise velocity plot is shown below. The plot shows that this configuration gives stronger and larger lateral structures starting at 1800m. The region of strong acceleration also remains stronger over more of the rotor plane.

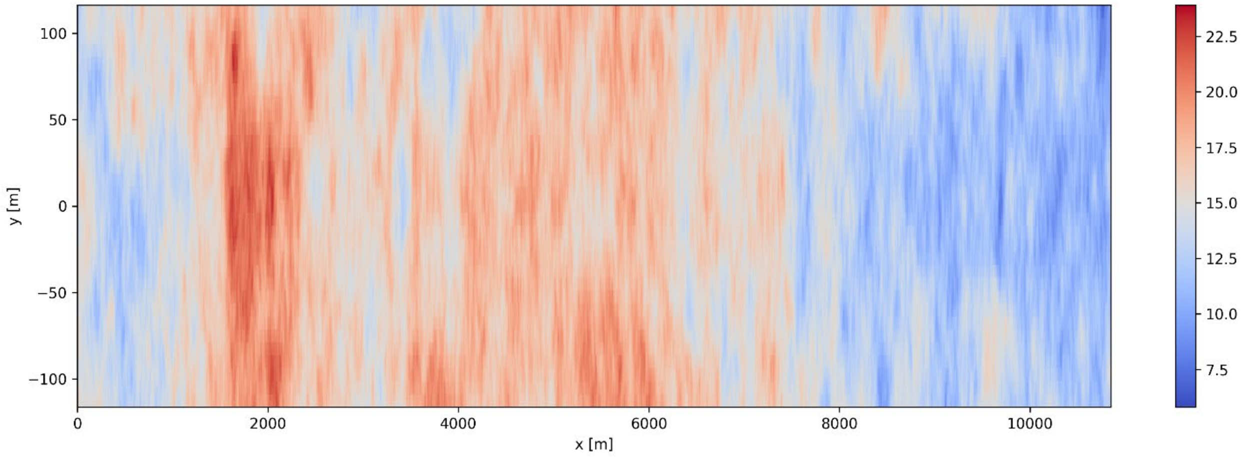


Figure 62 Stream-wise velocity [*m/s*] *at 100m for the coherent configuration*

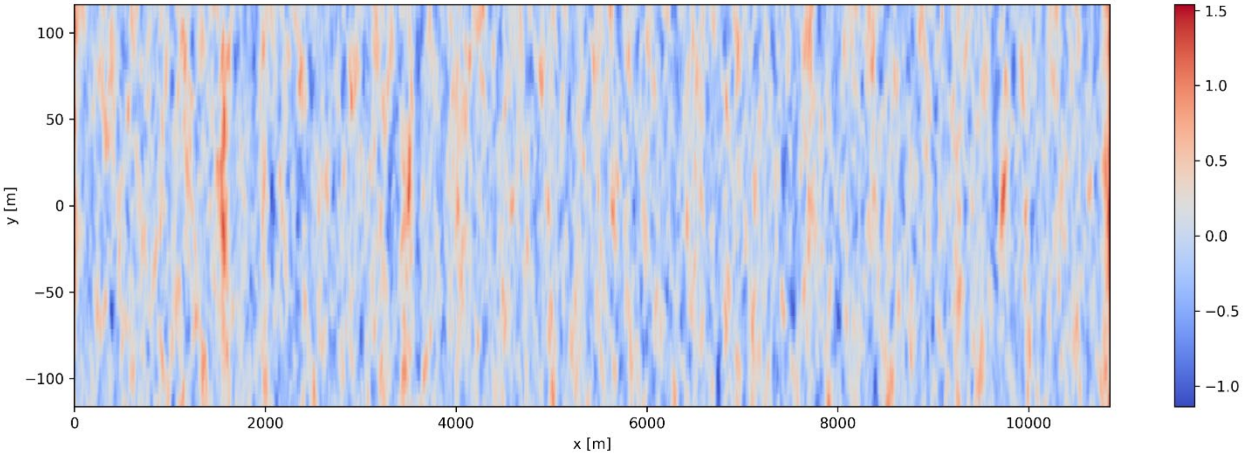


Figure 63 Filtered stream-wise acceleration $[m/s^2]$ *at 100m for the coherent configuration*

Finally, the next two plots show the stream-wise velocity for the side configuration along with the corresponding acceleration plot. Both the velocity and acceleration plot show how the side offset moves the turbulent structure laterally.

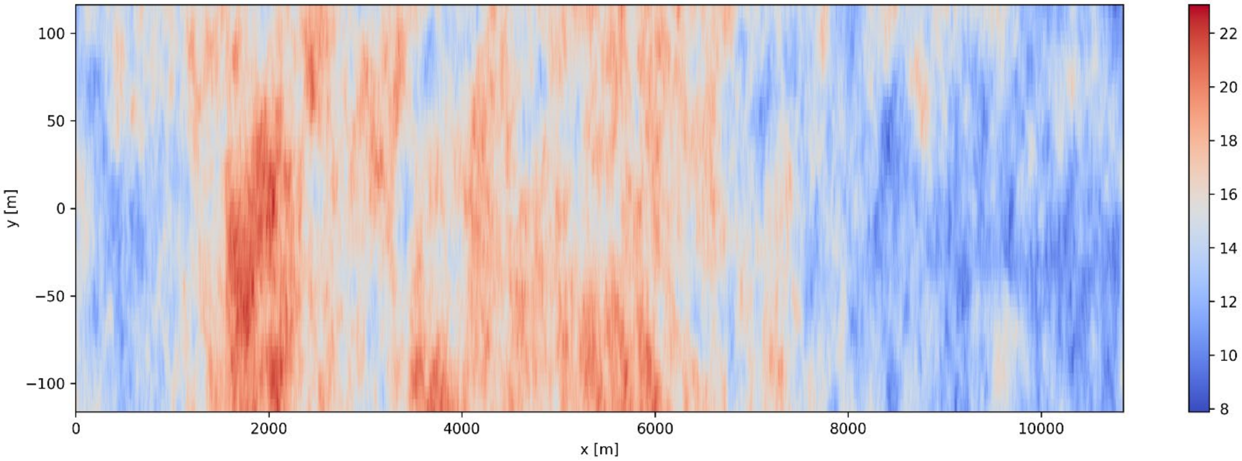


Figure 64 Stream-wise velocity [m/s] at 100m for the side configuration

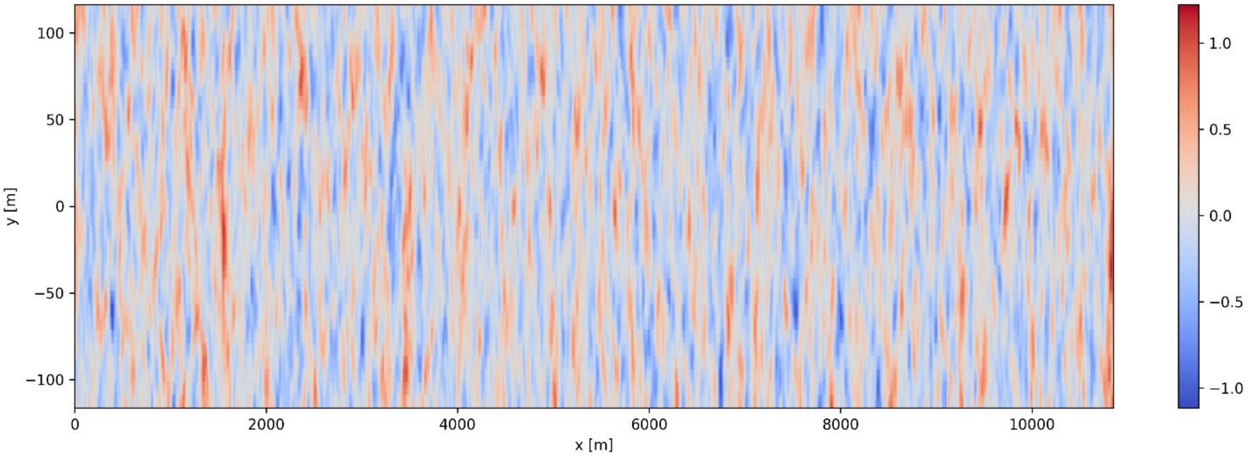


Figure 65 Filtered stream-wise acceleration $[m/s^2]$ *at 100m for the side configuration*

Summary of Loads due to Constrained Turbulence

The loads analysis focused on the tower bottom over-turning moment, tower top yaw moment, shaft main bearing bending moment about X, blade root flap-wise bending moment, tower-top acceleration in the Y direction and tip deflection.

The load statistics were processed in three ways:

1. **1% and/or 99% quantiles conditional on nominal wind speed**: All the loads, for each time step and all simulations in consideration with the same nominal wind speed are grouped together. Then based on this data, different quantiles of the data are presented. Since this analysis is aggregates data across simulations, a single simulation with exceptionally high loads can skew the data.
2. **10%, 50% and/or 90% quantiles of the maximum values of simulations conditional on nominal wind speed**: The minimum or maximum values (depending on whether the negative or positive values are significant) are collected from each simulation with a common nominal wind speed. Then the 10%, 50% and/or 90% quantile of these values is reported. This analysis would ignore the loads of any simulation that produced exceptionally high loads. However, since it is still based on the extremes, it can still be sensitive to stochastic effects.
3. **Overall cumulative distribution of loads over all simulations and wind speeds**: All the loads, for each time step and simulation are grouped together. Unlike the first method, the loads are not grouped by wind speed, these loads are also mixed together. Then these loads are presented as a cumulative distribution. These plots are used to investigate the overall effect of different simulation parameters on the loads.

An important consideration in interpreting the loads is the fact that the return period of the meteorology data is less than the assumed return period of the IEC extreme events. The IEC events have a target return period of 50 years, which represents the joint probability of atmospheric events, and for some load cases, the probability of simultaneous occurrence of an operational event such as a shutdown or a fault. Thus, it can be expected that the loads based on the constrained turbulence will increase once the loads are extrapolated to longer return periods or combined with operational events.

The loads analysis made the following comparisons

- Stream-wise acceleration events at 100m with the IEC extreme operating gust (EOG).
- Significant rate of direction change events at 100m with the IEC extreme direction change (EDC).
- Simulations with significant stream-wise acceleration and rate of direction change at any height, with the IEC extreme coherent gust and direction (ECD).

Each of these comparisons showed similar trends, thus for the sake of brevity, on the EOG loads are shown here. The other comparisons are given in the appendix.

Additional investigations looked at the effect of different parameters on the constrained loads simulations. The conclusions from this study are given afterwards referring to results presented in the same appendix.

Comparison of EOG Loads with Significant Stream-wise Acceleration Events

The IEC load simulations used the site-specific parameters for the turbulence and included shear. These simulations were compared against constrained turbulence simulations that were based on constraints applied with the larger turbulent length scale and added the wind shear in the load simulations. These simulations were selected since they were both most similar but are also realistic wind fields (*i.e.* including shear and background turbulence).

Below, Figure 66 and Figure 67 show the tower bottom bending moment. For the quantiles of conditional loads, it is clear that the constrained turbulence was able to give similar loads. It was only for the extremes above rated speed that the IEC loads significantly exceeded the constrained turbulence loads.

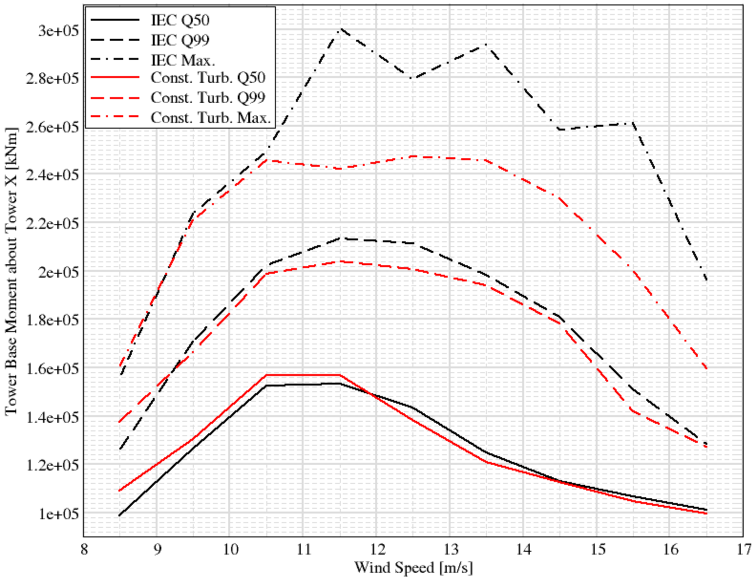


Figure 66 Quantiles of conditional tower bottom over-turning moment loads for IEC EOG and stream-wise acceleration at 100m simulations

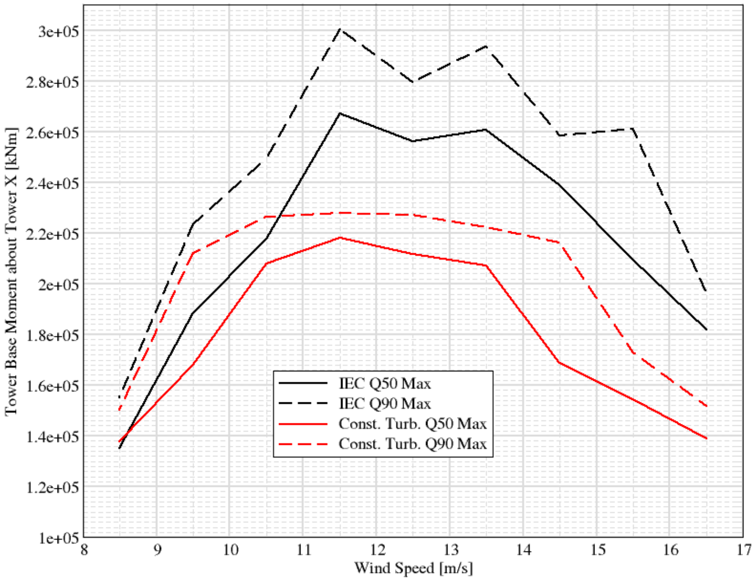


Figure 67 Quantiles of extreme loads for the tower bottom over-turning moment comparison between IEC EOG loads with stream-wise acceleration at 100m.

The tower-top yaw loads in the constrained turbulence simulations tend to meet or exceed the IEC loads below rated, while for above rated speed the IEC loads are greater (Figure 68, Figure 69).

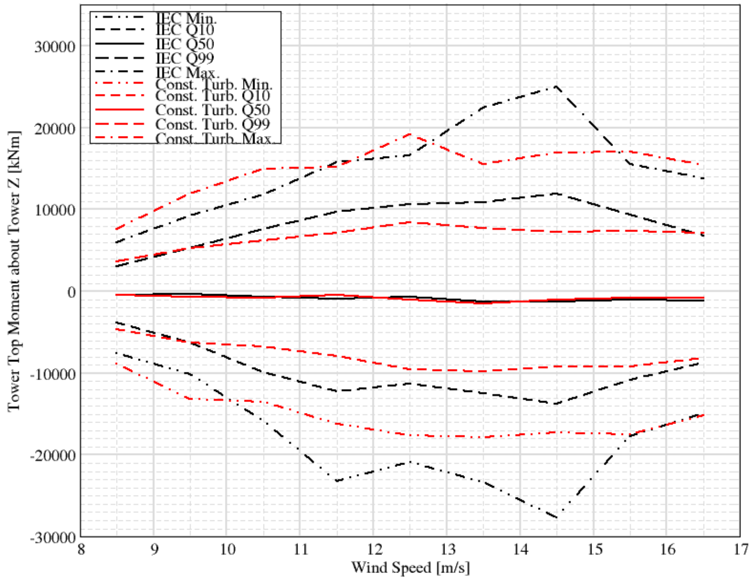


Figure 68 Quantiles of conditional tower top yaw moment for IEC EOG and stream-wise acceleration at 100m simulations

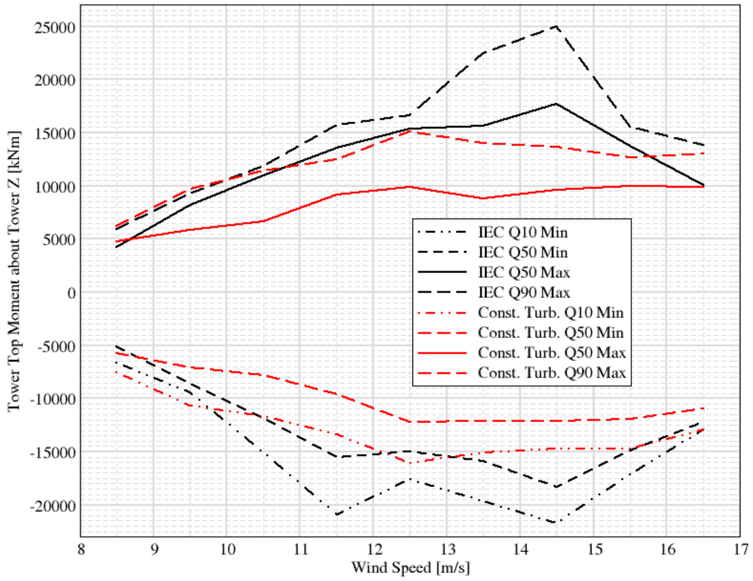


Figure 69 Quantiles of extreme loads for the tower top yaw moment comparison between IEC EOG loads with stream-wise acceleration at 100m.

The shaft main bearing bending loads (Figure 70) tend to be less in the constrained turbulence simulations. At 8.5m/s and 16.5m/s, the constrained turbulence simulations did produce similar loads.

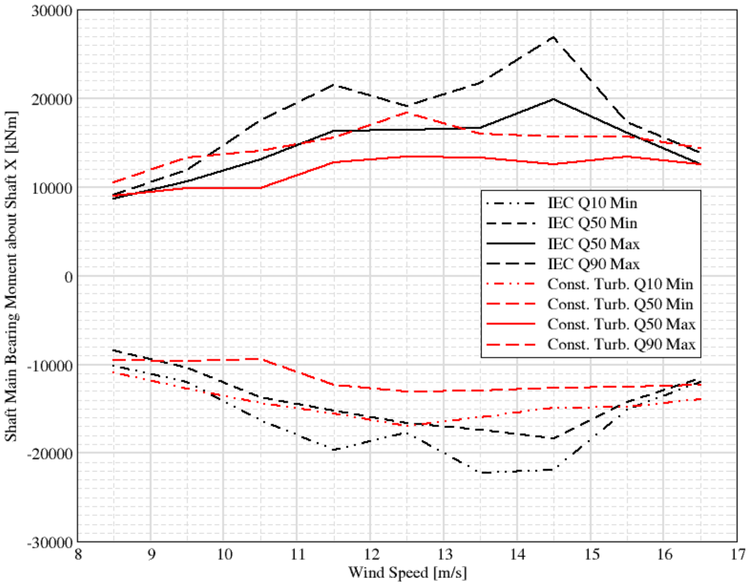


Figure 70 Quantiles of conditional shaft main bearing loads for the IEC EOG and significant stream-wise acceleration at 100m simulations.

The conclusions for the blade root flap-wise bending loads (Figure 71, Figure 72) are similar to those of the tower bottom over-turning bending moment. The quantiles of conditional loads are similar for both the constrained turbulence and the IEC loads. The absolute extremes for the constrained turbulence simulations have a greater absolute magnitude for all conditions. However, when looking at the quantiles of the extreme loads, this conclusion only holds for the 10% quantile below rated. Thus, there are 1-2 simulations with constrained turbulence that are responsible for very large extreme loads.

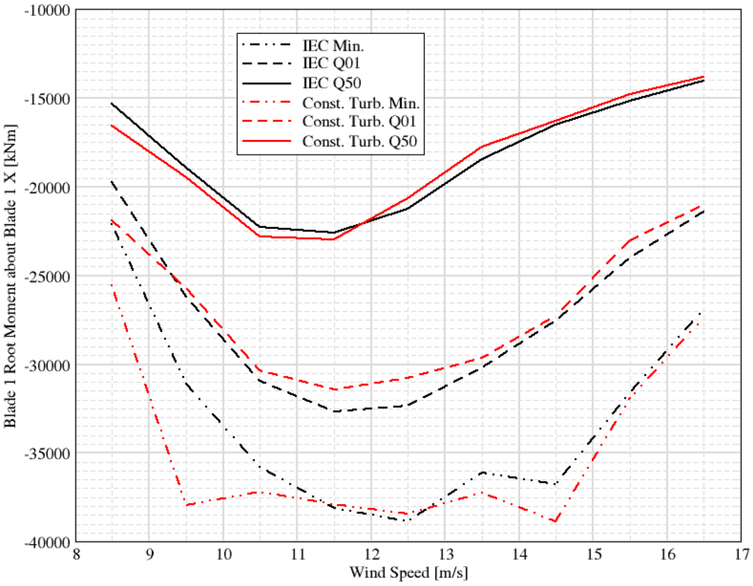


Figure 71 Quantiles of conditional loads for the root flap-wise bending moment; comparison between IEC EOG loads with stream-wise acceleration at 100m.

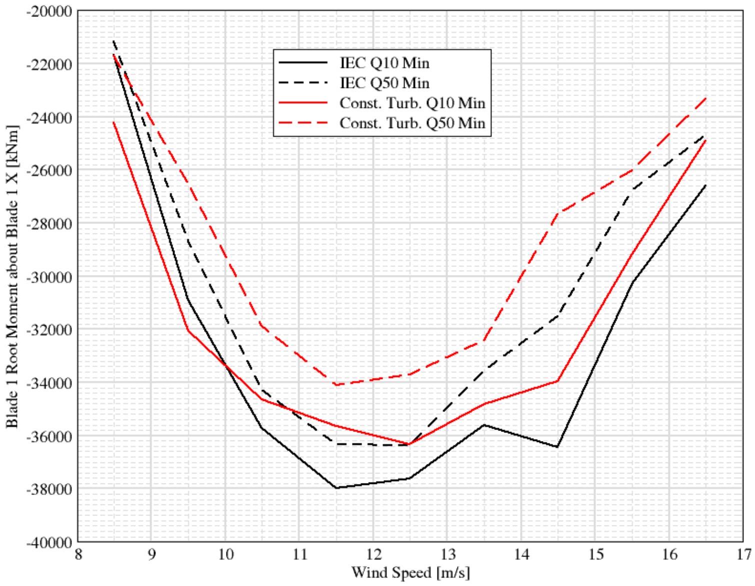


Figure 72 Quantiles of extreme loads for the root flap-wise bending moment; comparison between IEC EOG loads with stream-wise acceleration at 100m.

The constrained turbulence simulations produced the similar tower top acceleration loads over all wind speeds (Figure 73). Whereas the IEC EOG simulations produced significantly higher accelerations above rated speed. This suggests that the IEC EOG simulations are exciting tower vibrations or experiencing a larger impulse.

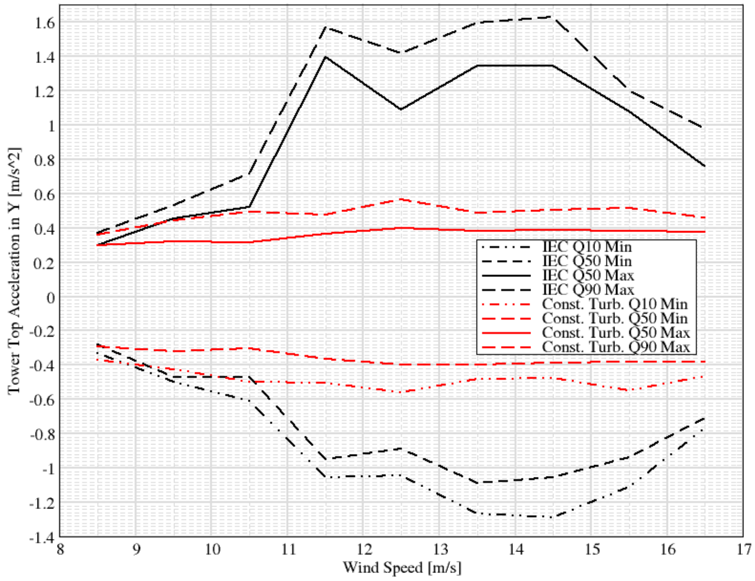


Figure 73 Quantiles of extreme loads for the tower top acceleration; comparison between IEC EOG loads with stream-wise acceleration at 100m.

The conclusions for tip-deflection (Figure 74, Figure 75) are similar to those for root flap-wise bending moments. The quantiles of conditional loads are similar between the constrained turbulence and IEC simulations. The extremes appear to be larger for the constrained turbulence. However, when looking at the

50% and 90% quantile of extremes, it appears that 1-2 simulations per wind speed are producing loads that exceed the IEC simulations.

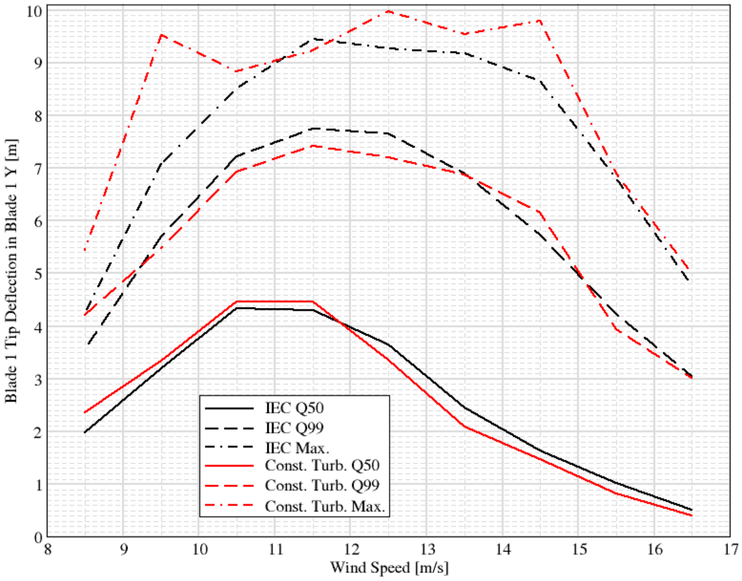


Figure 74 Quantiles of conditional blade tip deflection; Comparison between IEC EOG and significant stream-wise acceleration at 100m.

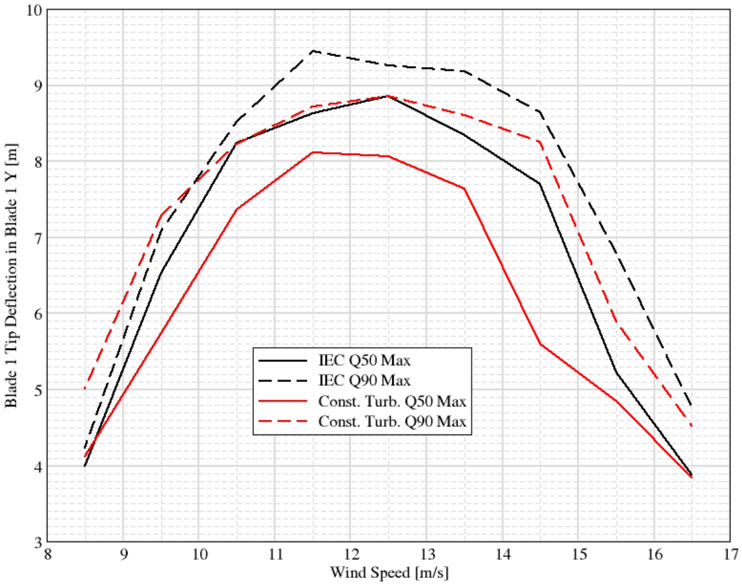


Figure 75 Quantiles of extreme tip deflection; Comparison between IEC EOG and significant stream-wise acceleration at 100m

The effect of different simulation parameters on the loads

Additional investigations looked at the effect of different parameters on the loads. For the sake of brevity, the charts are not given in the main report, instead they can be seen in the appendix. The main conclusions from these investigations are given here.

The constrained turbulence simulations looked at 5 different types of events: significant stream-wise acceleration at 100m and 160m, significant rate of direction change at 100m and 160m and then conditions with both a significant stream-wise acceleration and rate of direction change at any height. The loads from each of these events were compared to determine the types of events that gave the greatest loads. The cumulative distribution of loads showed that combined stream-wise acceleration and rate of direction change was greater than the other types of events. The number of simulations for the combined acceleration and rate of direction change is less than the other events, so the loads may not be statistically significant. These are followed by the other events, which all gave similar loads. Except for the significant rate of direction change events at 100m which produced slightly lower loads. The important outcome is that combined events produce higher loads, however the analysis did not collect many of these events, so these conclusions could be biased.

The spatial scale of the different events was investigated by using different length scales for applying the constraints and also by applying the constraints at different locations on the rotor disk. Load simulations where the constraints were applied with a smaller length scale tended to produce lower loads. This indicates the significant events attenuates quickly and is not seen significantly over the whole rotor disk. On the contrary, when the constraints are applied at 4 different locations over the rotor disk, with the large length scale, the loads were significantly higher. There is also evidence that some of the load simulations went into shutdown and produced some exceptionally high loads. However, the higher loads cannot be attributed to a minority of exceptional simulations. The quantiles of extreme loads also show that the large coherent event configuration produced larger loads. Thus, it is important to understand the spatial scale of these events to accurately determine the effect on loads.

The effect of shear in the constrained turbulence simulations was added in two ways. The first assumed that the application of the constraints would automatically include shear. The second subtracted a shear profile from the constraints and then added that shear to the load simulations. The second approach produced slightly higher loads. So, it appears that the shear effects in the constraints are attenuated too quickly when applied via constraints.

Additional studies were conducted on the IEC simulations to understand the effect of site-specific parameters. These parameters included the turbulence intensity, the parameters used in the Mann model and the recommended magnitude of the EOG and EDC based on the site-specific turbulence intensity. The results of these studies are not shown here or in the appendix, however, the results show that these parameters had negligible influence on the loads. This is not surprising considering that the estimated site-specific turbulence intensity was already similar to the IEC class 1C recommendation.

Furthermore, additional studies also looked at the effect of shear and a comparison between the normal turbulence model and extreme turbulence model. These studies showed that both shear and extreme turbulence intensity yield increased loads. These were the expected conclusions and thus these results are not shown here.

Summary of the outcome of the loads from constrained turbulence study

Overall, the loads obtained using the constrained turbulence were in most cases less than the loads calculated according to the IEC standards. However, there are also cases where the constrained turbulence loads were either similar or exceeded the IEC 61400 loads. Generally, the constrained turbulence loads tended to be stronger relatively below rated conditions. Another observation is that the highest IEC loads tended to be concentrated near rated wind conditions, while the highest constrained turbulence loads were more widely distributed over more wind speeds.

References in Section 6

[1] Mann, J. (1994). The spatial structure of neutral atmospheric surface-layer turbulence. Journal of Fluid Mechanics, 273, 141–168. https://doi.org/10.1017/S0022112094001886

[2] Mann, J. (1998). Wind field simulation. Probabilistic Engineering Mechanics, 13(4), 269–282. https://doi.org/10.1016/s0266-8920(97)00036-2

[3] Dimitrov, N. K., & Natarajan, A. (2017). Application of simulated lidar scanning patterns to constrained Gaussian turbulence fields for load validation. Wind Energy, 20(1), 79–95. https://doi.org/10.1002/we.1992

[4] Hipersim Manual, https://gitlab.windenergy.dtu.dk/HiperSim/hipersim, accessed March 2022.

[5] Natarajan, A., Yde, A., Bak, C., Zahle, F., Blasques, J. P. A. A., Henriksen, L. C., … Kim, T. (2013). The DTU 10-MW Reference Wind Turbine.

[6] IEC (2005) International Standard IEC61400-1: wind turbines—part 1: design guidelines, Third edition.

[7] Larsen, T. J., & Hansen, A. M. (2007). How 2 HAWC2, the user's manual. Risø National Laboratory.

[8] Hiperwind Deliverable 2.3: Environmental joint probability distributions and uncertainties. https://www.hiperwind.eu/publications (accessed March 2022).

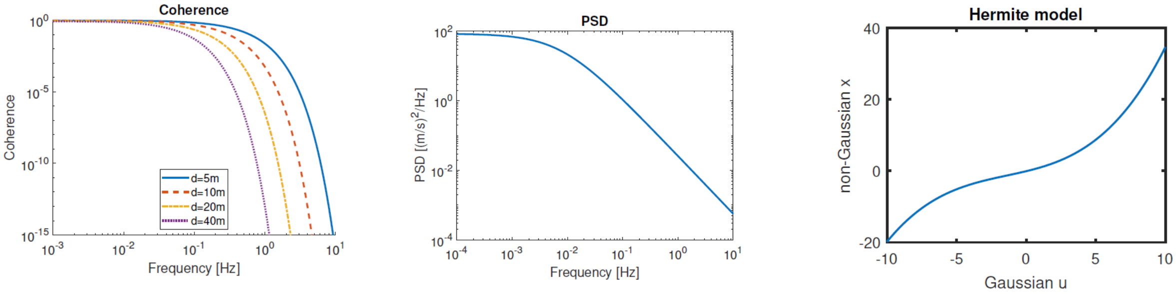
7. Non-Gaussian turbulence time series generation for load simulations

7.1. Non-Gaussian turbulence box generation based on target spectra and statistical moments

Several approaches were studied for the purpose of generation of non-Gaussian turbulence box.

One of the approaches is based on a combination of the Hermite model and the correction of the spectrum. It is essentially a purely statistical approach that seeks to match the prescribed values of skewness and kurtosis while keeping the properties of coherence and spectrum (Figure 76).

This statistical approach is implemented by using Veer's method [1]. In Veer's method, the Kaimal spectrum model is used for all the three wind velocity components, while an exponential coherence model is imposed on the longitudinal wind velocity component. In future work, the coherence model and spectrum model can also be based on the Mann model [2] or LES simulation data.



(a) Coherence model (b) Kaimal spectrum (c) Hermite model

Figure 76 Generation of non-Gaussian turbulence box using a statistical approach

One key step in Veer's method is the Cholesky decomposition, which is used when imposing spatial coherence. Two numerical schemes were studied concerning the Cholesky decomposition. In one version, the Cholesky decomposition is directly applied to the coherence matrix. In the other version, the Cholesky decomposition is applied to an approximation of the coherence matrix by ignoring the small numbers within a given tolerance. This approximation leads to sparse matrices for a range of frequencies and thus provides a practical scheme to speed up the computation.

After generation of the Gaussian turbulence box, the Hermite model (Winterstein, 1988 [3]) is used to map the Gaussian turbulence box to a non-Gaussian turbulence box with prescribed skewness and kurtosis. The basic idea underlying the Hermite model is to transform a Gaussian distribution into a non-Gaussian distribution. However, the Hermite model not only changes the statistical distribution but also changes the spectrum.

In order to correct the spectrum, the correlation-distortion approach (Gurley et al., 1997 [4]) is used in this study. The spectrum is first modified such that after applying the Hermite model to the Gaussian turbulence box the desirable spectrum is obtained. In the numerical implementation, this modified spectrum is computed in a few steps, utilizing the property that the point spectrum and the autocorrelation function are Fourier transform pairs. First, the undistorted correlation function is obtained from the theoretical spectrum via an iFFT. Second, the correlation function is distorted by solving the nonlinear algebraic equations defined in [4]. Finally, the distorted correlation function is transformed back to a distorted spectrum via FFT. The procedure is illustrated in Figure 77.

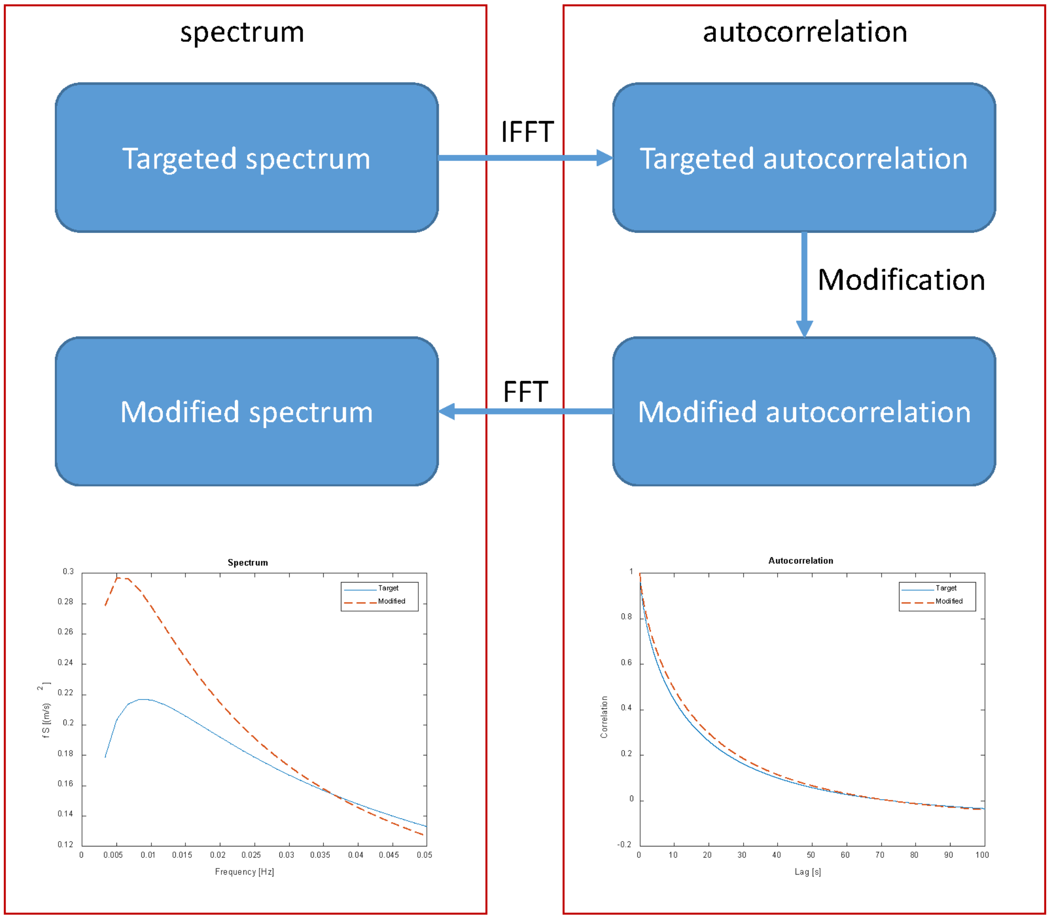


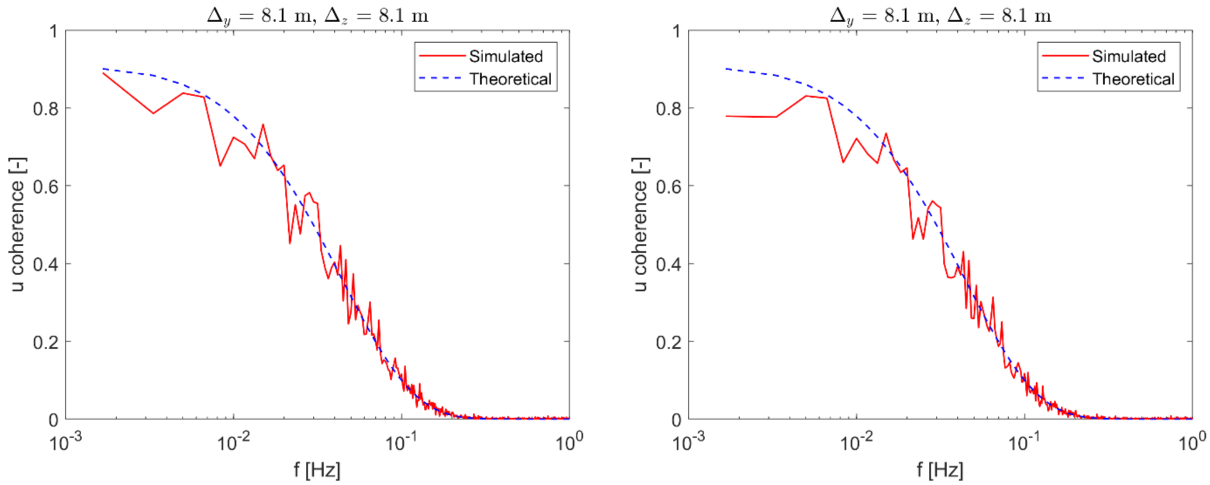
Figure 77 Procedure for spectrum adjustment to ensure the target properties are attained after applying a Hermite transformation.

The above statistical approach is applied to generate a number of non-Gaussian turbulence boxes with skewness in [-1,0] and kurtosis in [3,6]. Representative results are shown in the figure below for a combination of skewness=-1 and kurtosis=6. The other parameters are shown in Table 18 below.

Table 18: Parameters for generating the results of Gaussian and non-Gaussian turbulence boxes

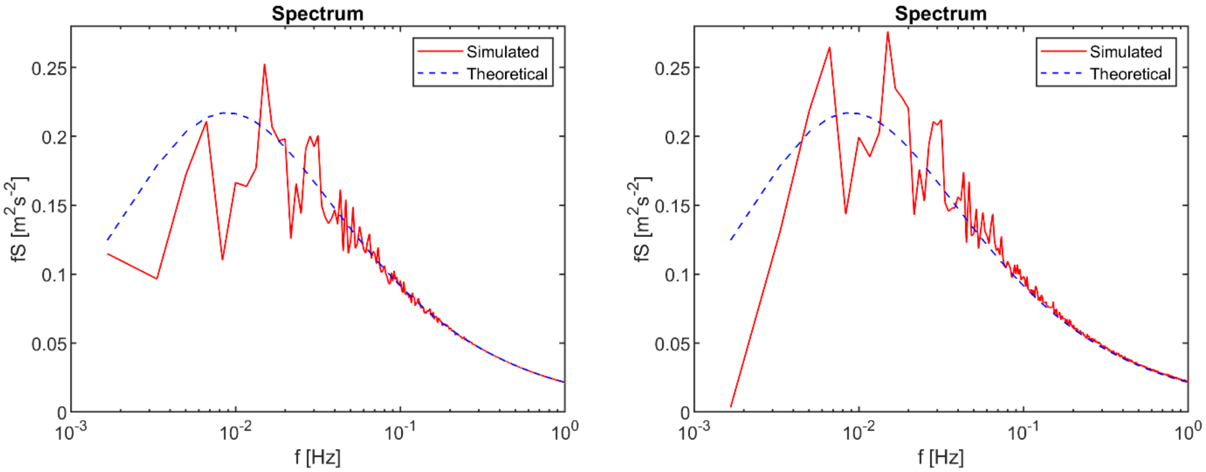
parameters	**values**
Uhub	12 m/s
SigmaU	1
Lu	340.2
Nx	8192
Ny	32
Nz	32
Dx	0.879 m
Dy	8.129 m
Dz	8.129 m
seed	0

In the following Figure 77, the simulated coherence and spectrum are estimated by using many time series in one turbulence box. The non-Gaussian turbulence box has coherence and spectrum that are similar to the Gaussian turbulence box. The major difference lies in the lowest frequency in the turbulence box. The non-Gaussian turbulence box has a lower coherence and less energy in the lowest frequency compared to the Gaussian turbulence box.



(a) Coherence of Gaussian turbulence box

(b) Coherence of non-Gaussian turbulence box

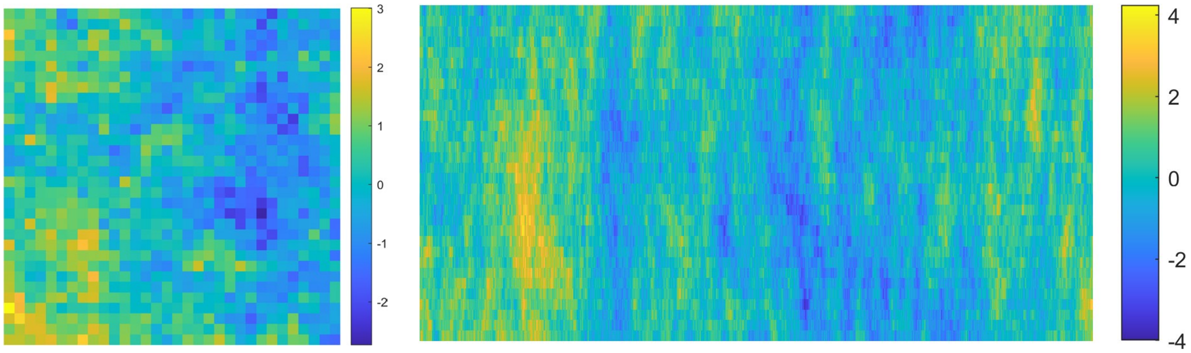


(c) Spectrum of Gaussian turbulence box

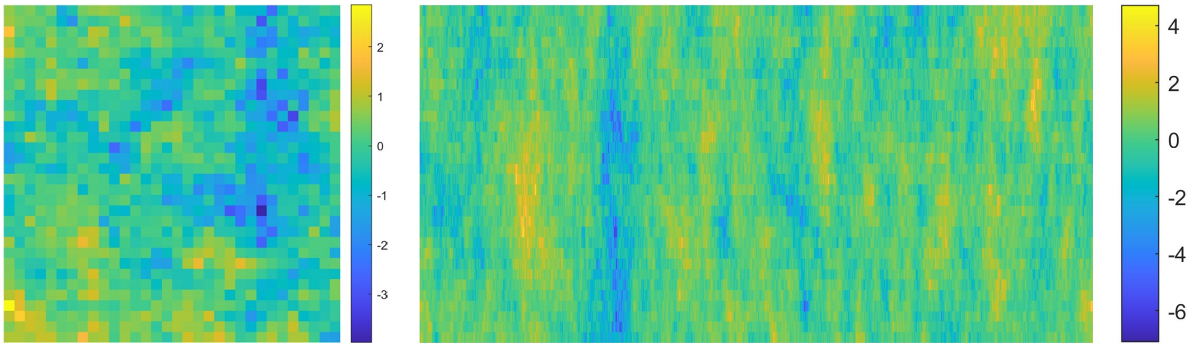
(d) Spectrum of non-Gaussian turbulence box

Figure 78 Comparison of the properties of coherence and spectrum for Gaussian and non-Gaussian turbulence boxes generated by using the statistical approach

Figure 78 compares the YZ-plane and XZ-plane of Gaussian and non-Gaussian turbulence boxes. It can be seen that in non-Gaussian turbulence box, the area of bright color is much larger than that in Gaussian turbulence box due to skewness. However, the minimum value (i.e., –6 m/s) in non-Gaussian turbulence box is less than that in Gaussian realization (–4 m/s). This implies that non-Gaussian turbulence box can present a more drastic change in wind velocity than Gaussian turbulence box.



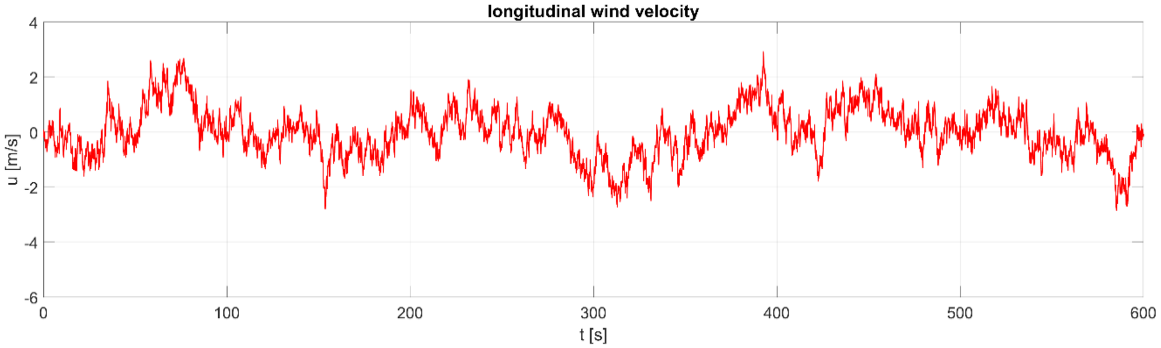
(a) YZ-plane of Gaussian turbulence box (b) XZ-plane of Gaussian turbulence box



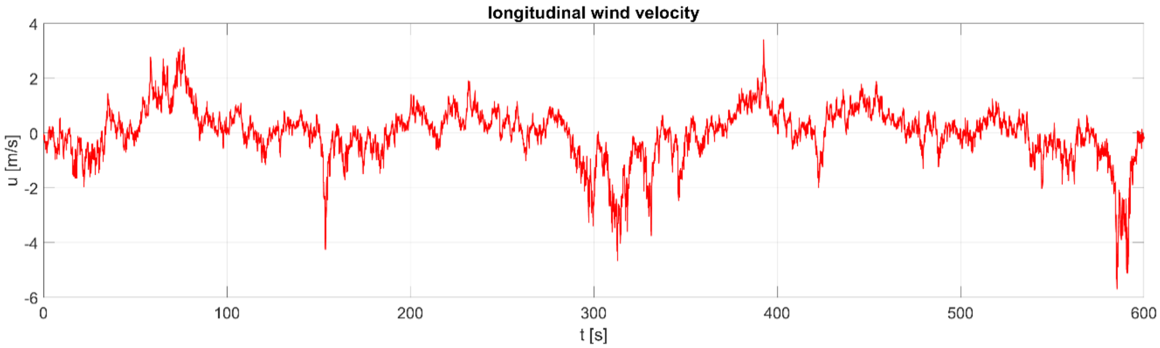
(c) YZ-plane of non-Gaussian turbulence box (d) XZ-plane of non-Gaussian turbulence box

Figure 79 Comparison of YZ- and XZ-planes of Gaussian and non-Gaussian turbulence boxes

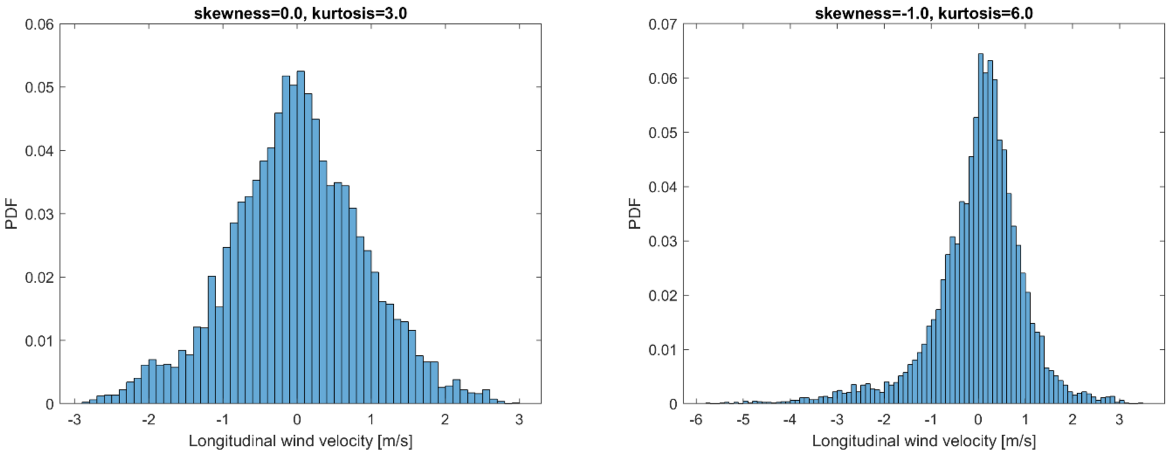
Figure 79 compares the time series in Gaussian and non-Gaussian turbulence box. It can be clearly seen that the non-Gaussian time series has a larger fluctuation than the Gaussian time series. This larger fluctuation may impact the dynamic loads and consequently the fatigue of the wind turbine components.



(a) One time series in Gaussian turbulence box



(b) One time series in non-Gaussian turbulence box (skewness=-1 and kurtosis=6)



(c) Histogram of Gaussian time series (d) Histogram of non-Gaussian time series

Figure 80 Comparison of time series in Gaussian and non-Gaussian turbulence boxes

7.2. Impact of non-Gaussian turbulence boxes on turbine loads

The Veers spectral model with non-Gaussian statistics was used to evaluate the impact of non-Gaussian statistics on the loads. The simulations varied skewness between 0 and -1, then kurtosis between 3 to 6. Furthermore, the standard deviations in the stream-wise component was according to the normal turbulence model and the extreme turbulence model.

The figure below shows one example of the turbulence realization in the stream-wise component. The nominal wind speed was 18m/s. This example was chosen because it exhibited the largest fluctuation, thus all other fluctuations are smaller. What can be seen is that the negative skewness biases the larger fluctuation to smaller values, while the kurtosis amplifies the magnitude of the larger fluctuations.

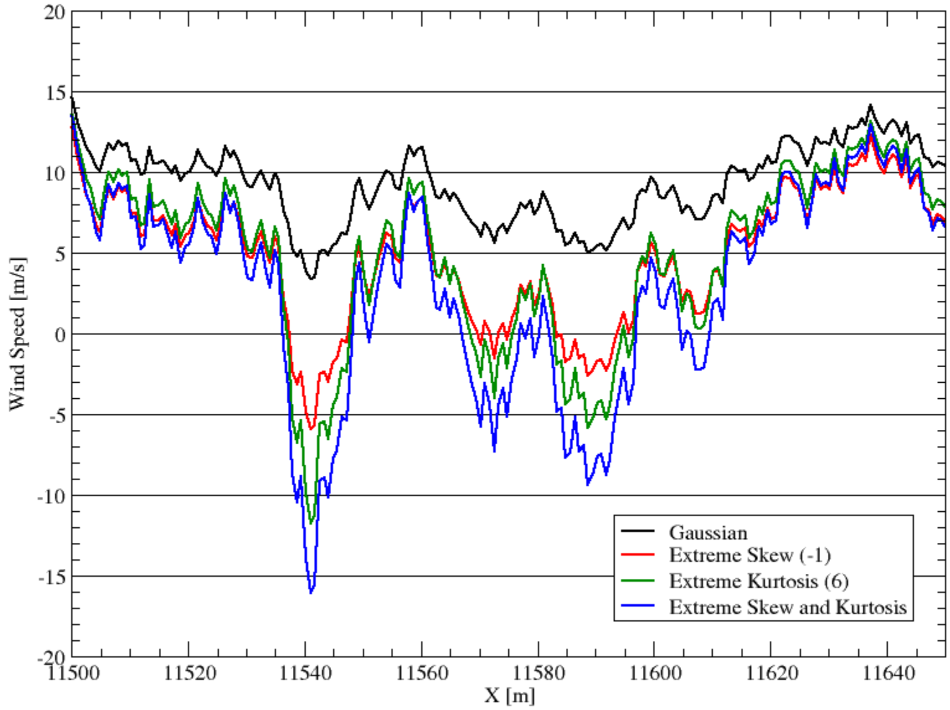
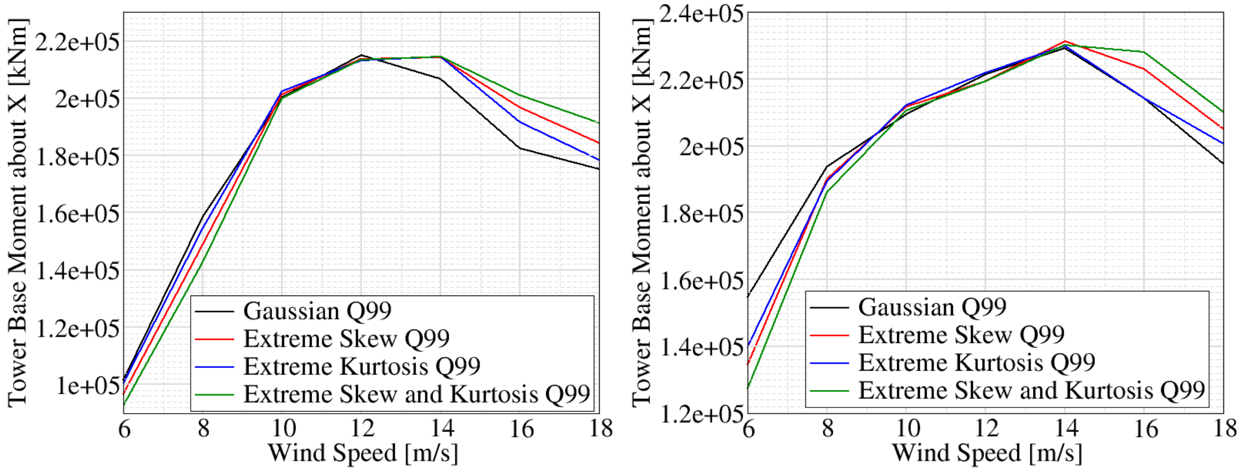


Figure 81 Examples of Non-Gaussian Turbulent Fluctuations

It is unlikely that the fluctuation in the figure above is physical. The most extreme case was selected to better show the effect of adding skewness and kurtosis. So, what is shown here may be an outlier, but at the same time care is needed to avoid conclusions based on unphysical conditions. In this work we did not filter out the unphysical conditions, so the conclusions here are based on the effect of skewness and kurtosis as applied here.

A variety of load simulations, with the same 6 seeds was used for all points in the design of experiments. The loads simulation was based on the DTU 10MW reference wind turbine [5] that has already been described above. The loads statistics presented here are based on the 1% and 99% quantile of all the values at all time steps and simulations. The statistics based on the absolute maximum or minimum was too noisy to draw conclusions.

The following figures show the tower bottom bending moment for both turbulence models. The effect of the non-Gaussian statistics does not show significant effect on the absolute maximum of the loads, instead it shifts the loads from below rated conditions to above rated. As a result, the tower-bottom loads under non-Gaussian conditions are lower for wind speeds below rated, and higher for wind speeds above rated. This is likely due to the negative skewness: the kurtosis does not show a significant impact. The negative skewness causes the instantaneous wind to reach rated conditions more frequently, where loads are typically highest. The increased turbulence intensity in the ETM case leads to higher loads than with NTM regardless of other inputs, while the conclusions on Non-Gaussian statistics hold for both the ETM and NTM cases.

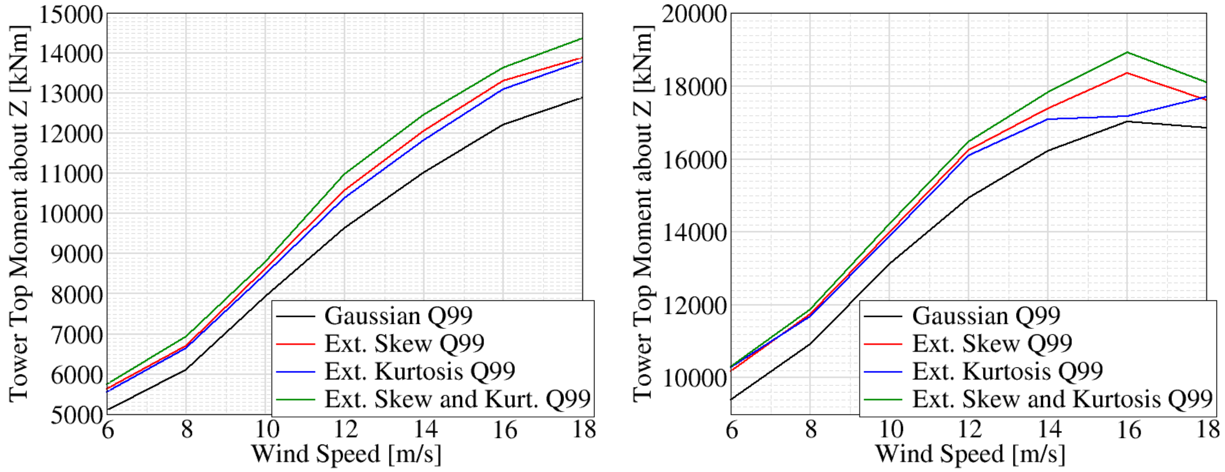


Tower Bottom Moment with NTM

Tower Bottom Moment with ETM

Figure 82 Effect of wind field non-Gaussianity on the tower base fore-aft bending loads

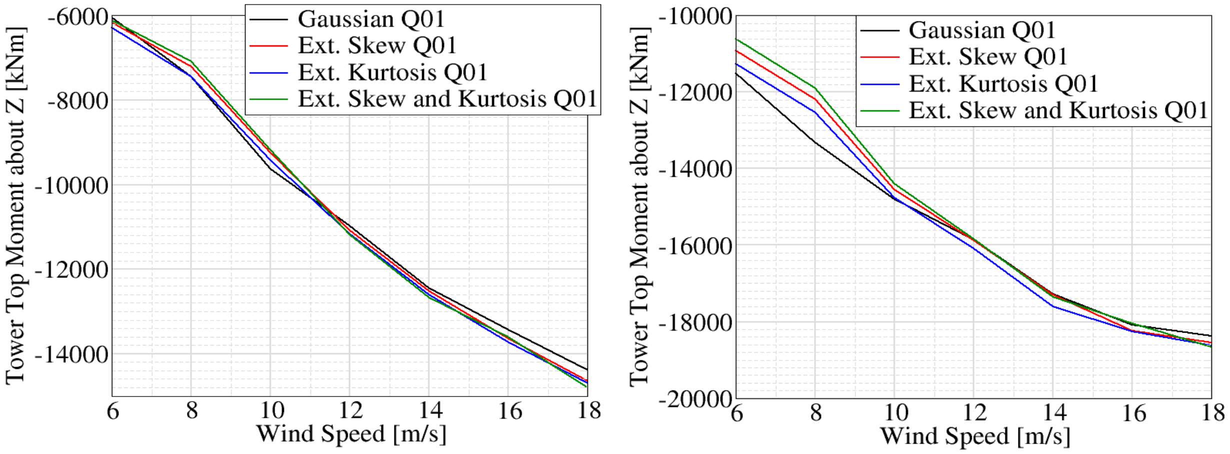
The following figures show the tower top yaw moment. The absolute magnitude of this moment was similar for both positive and negative values, so the charts show both the 99% and 1% quantile. The non-Gaussian statistics do appear to have a significant effect of increasing these yaw loads, while for the 1% quantile they do not have a significant impact. This shows that there is an asymmetric effect of these statistics on the rotor.



Tower Top Yaw Moment Q99 with NTM

Tower Top Yaw Moment Q99 with ETM

Figure 83 Effect of wind field non-Gaussianity on the tower top yaw moment maxima

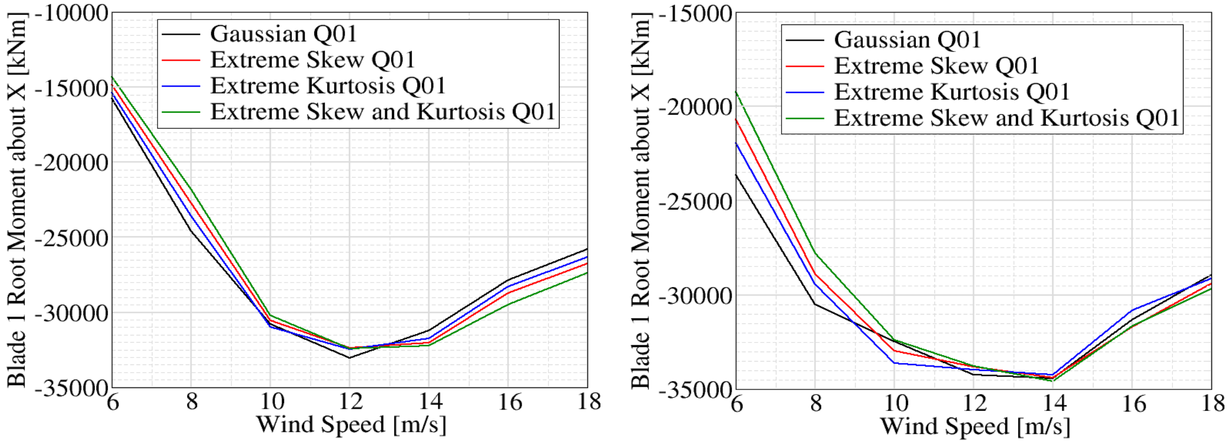


Tower Top Yaw Moment Q01 with NTM

Tower Top Yaw Moment Q01 with ETM

Figure 84 Effect of wind field non-Gaussianity on the tower top yaw moment minima

Finally, the following plots show the blade root flap-wise moment. The effect of non-Gaussian statistics is similar here as with the tower bottom bending moment.



Blade Root Flap-wise Moment with NTM

Blade Root Flap-wise Moment with ETM

Figure 85 Effect of wind field non-Gaussianity on the tower base fore-aft bending loads

References in Section 7

[1] Veers. P. (1988) Three-Dimensional Wind Field Simulation. SANDIA report SAND88-0152

[2] Mann, J. (1994). The spatial structure of neutral atmospheric surface-layer turbulence. Journal of Fluid Mechanics, 273, 141–168. https://doi.org/10.1017/S0022112094001886

[3] Winterstein, S. R. (1988). Nonlinear vibration models for extremes and fatigue. Journal of Engineering Mechanics, 114(10), 1772-1790.

[4] Gurley, K. R., Tognarelli, M. A., & Kareem, A. (1997). Analysis and simulation tools for wind engineering. Probabilistic Engineering Mechanics, 12(1), 9-31.

[5] Natarajan, A., Yde, A., Bak, C., Zahle, F., Blasques, J. P. A. A., Henriksen, L. C., … Kim, T. (2013). The DTU 10-MW Reference Wind Turbine.

[6] Dimitrov, N. K., & Natarajan, A. (2017). Application of simulated lidar scanning patterns to constrained Gaussian turbulence fields for load validation. Wind Energy, 20(1), 79–95.

8. Conclusions

The present report studied two aspects of the inflow of wind turbines in wind farms that are a major factor affecting the wind turbine loads. The first aspect is wake effects and more specifically how to describe (parameterize) them in terms of quantities suitable for surrogate model inputs – and what is the sensitivity of the loads to these inputs. We studied two parameterization methods – one based on the wind farm layout such as relative turbine positions, and another one based on using shape functions to describe the wake deficit properties. These methods were tested with surrogate models trained on aeroelastic load simulations under waked conditions, with the following general conclusions:

- It is feasible to parameterize the wind farm layout in a way that enables the construction of surrogate models that predict loads and power outputs of individual turbines in arbitrary wind farm layouts;
- The sensitivity analysis showed that the ambient wind speed is the governing factor for loads and power production. The wind direction (which determines the strength of the wake effects) and the ambient turbulence are also of high importance. For wind speeds just below rated where the turbine thrust is highest, the wind direction has higher importance than the turbulence, while at low wind speeds the turbulence has more significant effect. For all methods that were studied, the wind shear was found to be of small significance.
- The performance of the surrogates is affected by the significance of wake effects: for signals with less wake impact, such as the extremes, the performance is consistently higher

The second inflow aspect we considered is the turbulent field that is used as input to load simulations. The focus is on situations that deviate from the normal operation under stationary conditions where the standard turbulence generation approaches provide sufficient quality of results. The special scenarios considered were transient events (extreme wind gusts, ramps and direction changes), and situations with highly non-Gaussian field statistics. For the purpose of this investigation, an open Python-based turbulence field generation tool was created within the Hipersim software package. The tool generates turbulence fields that can serve as inputs to aeroelastic load simulations, and can embed measured wind time series as constraints in the wind field. The capabilities were further expanded by introducing the possibility to generate non-Gaussian fields with a predefined skewness and kurtosis.

Using the constrained simulation approach, the load response of the DTU10MW turbine was simulated under a set of transient events that were previously obtained from measurements as part of Hiperwind Deliverable 2.3. The load results from these events were compared to simulations with the transient design load cases defined in the IEC61400-1 standard. In addition, the effect of introducing non-Gaussian statistics was assessed by carrying out pairs of load simulations with the same turbulence seeds but different skewness and kurtosis. The following conclusions were drawn:

- It was verified that the Hipersim/Turbgen tool produces turbulence boxes with the correct spectrum and coherence properties as prescribed with the Mann turbulence spectrum;
- The constrained simulation functionality of Hipersim results in correct reproduction of the target time series in the turbulence boxes;
- Overall, the loads obtained using simulations with measured transient events were in most cases less than the loads resulting from the synthetic events prescribed by the IEC standards. However, there are also cases where the constrained turbulence loads were either similar or exceeded the IEC 61400 loads;
- The synthetic conditions from the IEC standard tend to produce a different relationship between wind speed and load magnitude, compared to the measured time series simulations. The highest

loads from IEC events tended to be concentrated near rated wind conditions, while the highest constrained turbulence loads were more widely distributed over more wind speeds;

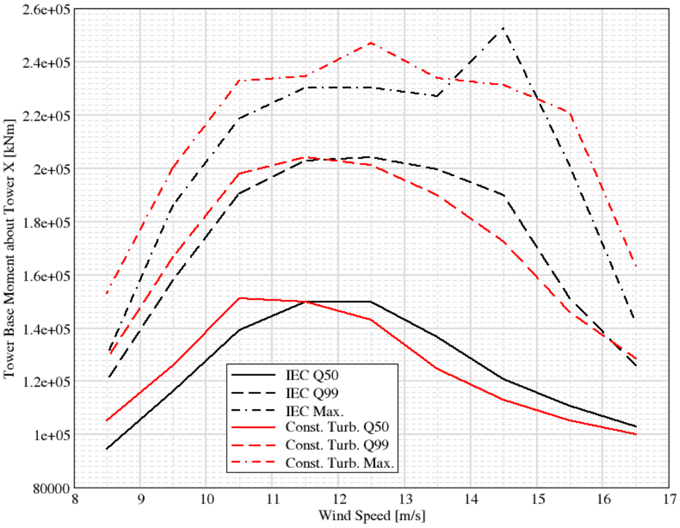
- A non-Gaussian turbulence box generator was developed that produces turbulence boxes with the Veers model and arbitrary skewness and kurtosis. It was verified that the resulting turbulence fields attain the target statistical moments, while retaining the correct spectrum and coherence;
- The present method of non-Gaussian turbulence generation is based on statistical rather than physical considerations. Although the results are statistically consistent, there is a chance that they are unphysical as e.g. major intermittent wind direction reversals. Increasing the skewness and kurtosis of a wind time series amplifies the chance of such non-physical events;
- For most load channels, the effect of non-Gaussian wind statistics was mild, such as minor changes in the design driving load and a shift in the wind speed where the dominant load events occur. An exception were the tower top yaw moment extremes, where changes in the skewness and kurtosis led to a significant increase in the loads.

Appendix A – Additional Results from the Loads Evaluation of Constrained Turbulence

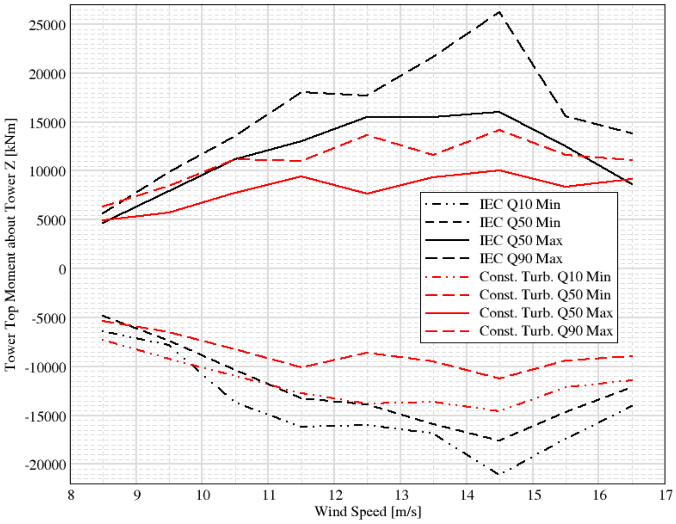
The following set of plots and points are taken from the comparison between IEC loads simulations and constrained turbulence simulations. The appendix gives the observations in point form, followed by the plot supporting the observations.

Comparison of EDC Loads with significant rate of direction change events at 100m

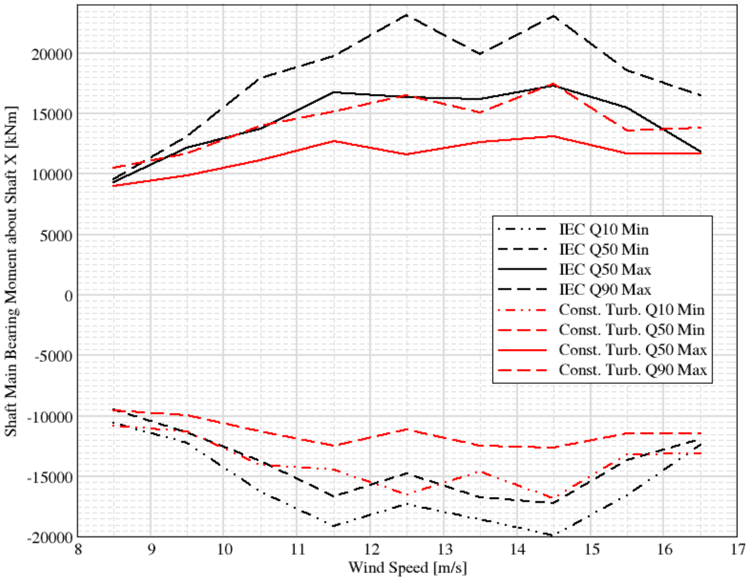
- The source for this is: “edc vs phi100 IEC-MS-TD-GD-SS-edc EVT-SS-CL-phi100”
- Tower bottom bending moment
    - The Q99 show it best, but basically the constrained loads are higher at/below rated



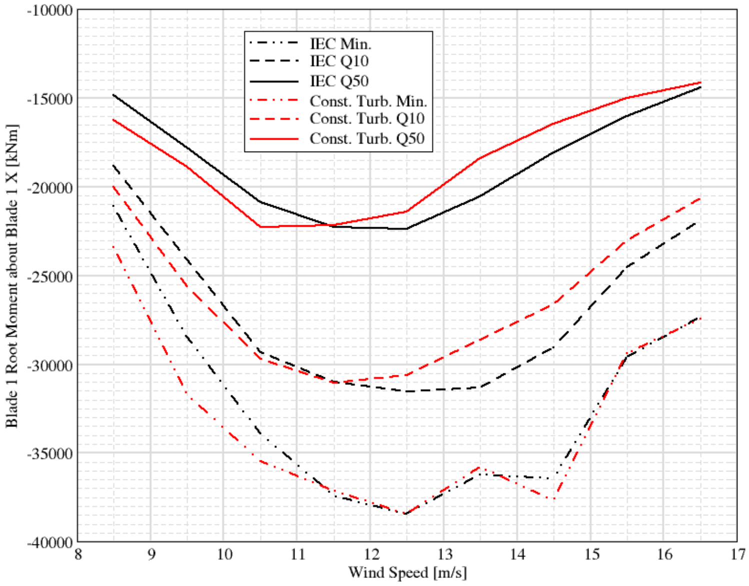
- Tower top yaw loads
    - Overall the constrained loads are lower … but at the smallest wind speed the constrained loads appear to exceed.

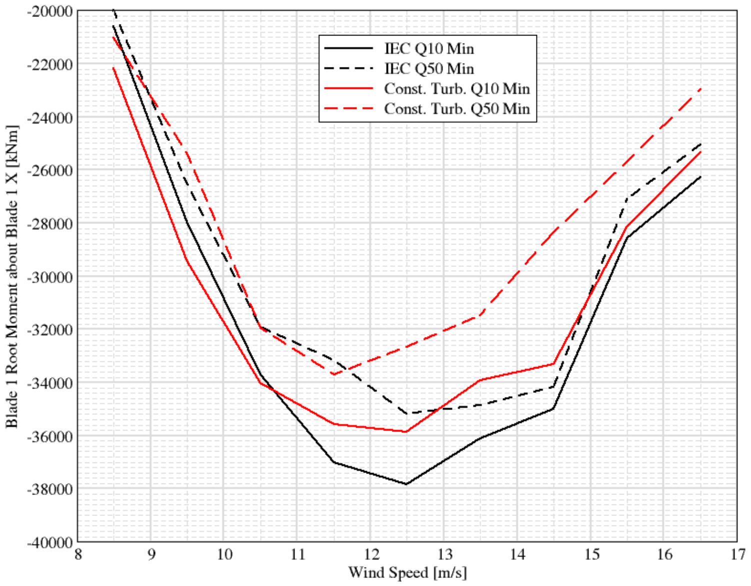


- Shaft Main Bearing Moment XY
    - Generally, the constrained turbulence loads are less, but at 8.5, they are nearly the same and sometimes slightly larger

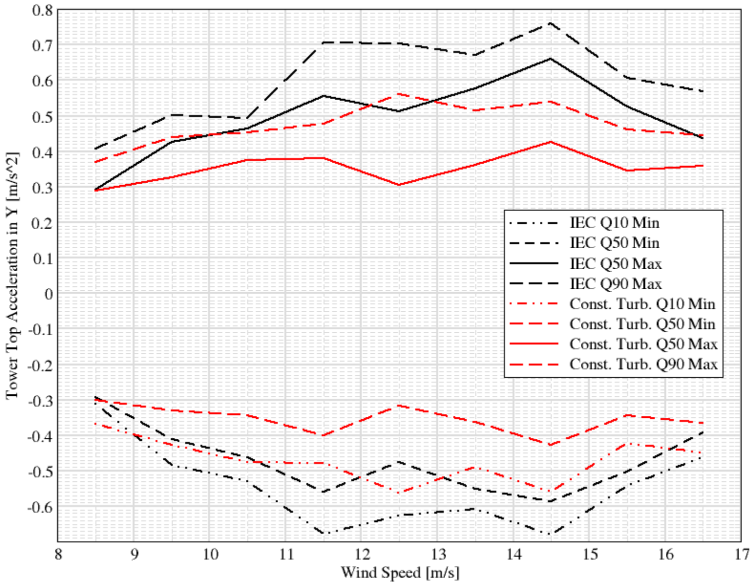


- Blade root flap-wise bending moment
    - For the Q99 loads, the constrained turbulence exceeded the IEC loads at or below rated.
    - However for the stats of extremes, that was also the case
    - The absolute max is the same for both

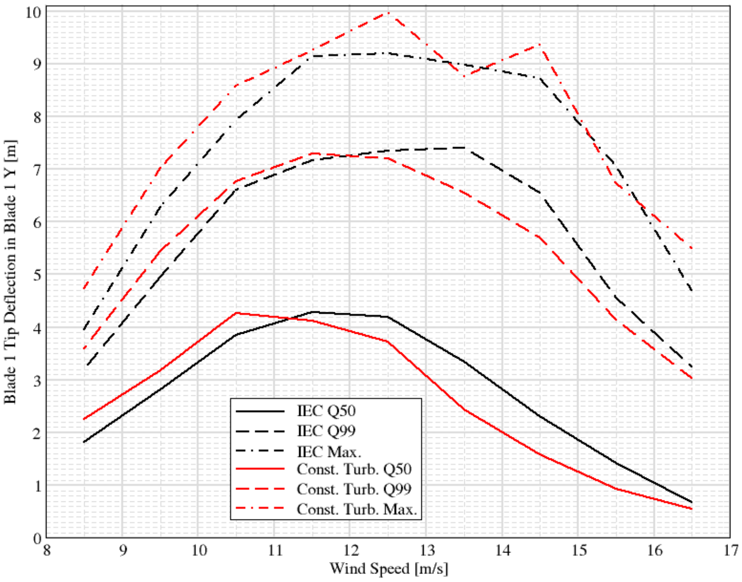




- Tower top acceleration Y
  - Generally the constrained loads are less

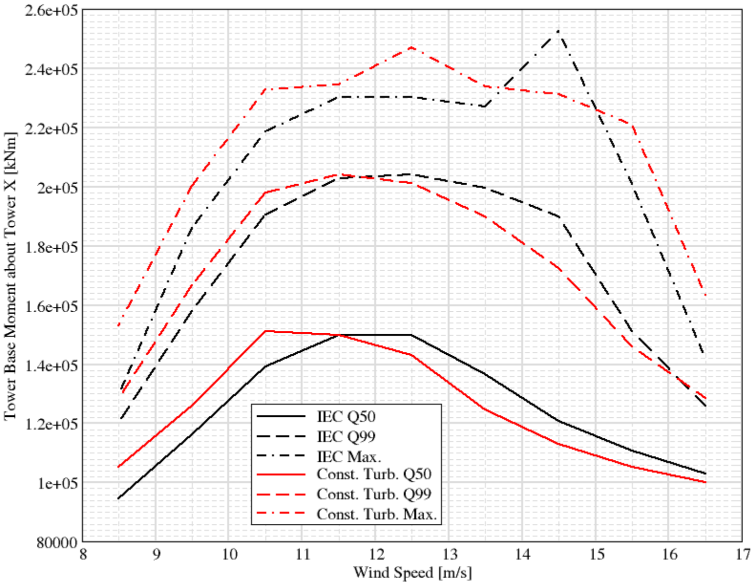


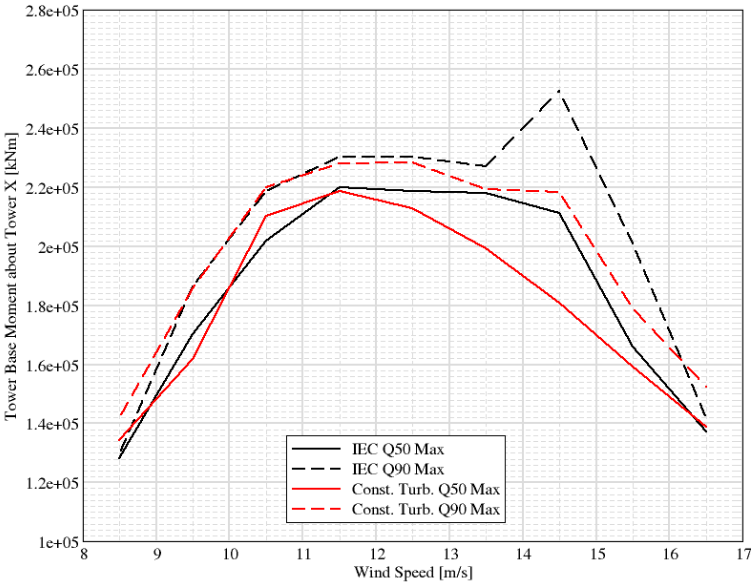
- Tip deflection
  - Constrained loads are greater at or below rated, then the IEC loads are greater
  - The absolute maximum value is the same for both



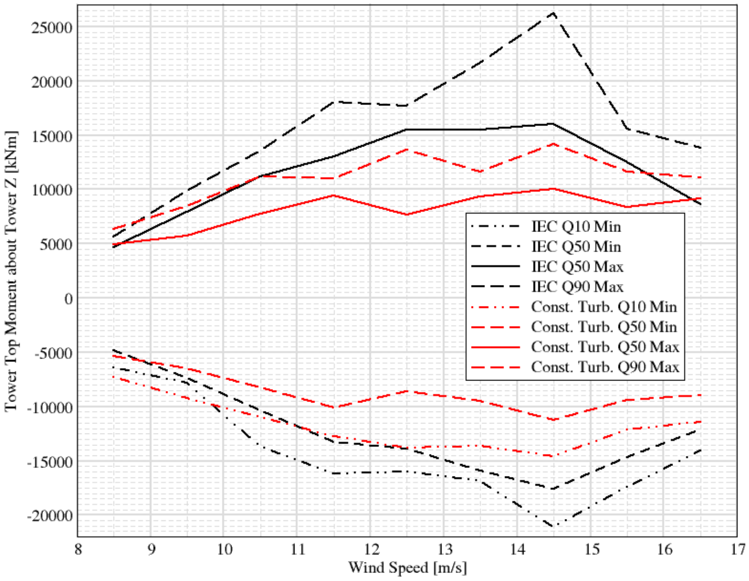
Comparison of ECD Loads with significant stream-wise acceleration and rate of direction change in constrained turbulence at any height simulations

- An overall point is that these conclusions are based on a small subset of all the simulations
- The source for these observations are from " ecd vs auPhi100160 IEC-MS-TD-GI-SS-ecd EVT-SS-CL-auPhi"
- Tower bottom bending moment
    - So the constrained loads appear to be larger for the Q99 over all wind speeds
    - However the IEC loads are higher for the statistics of the max (with some exception)

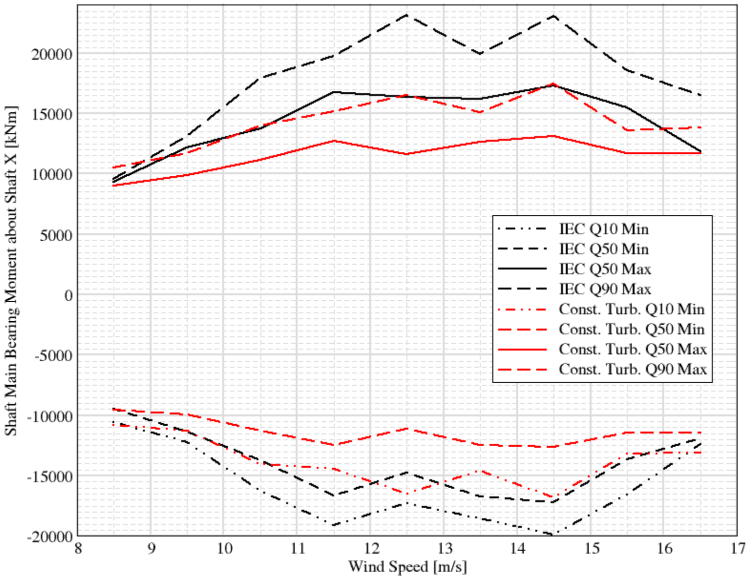




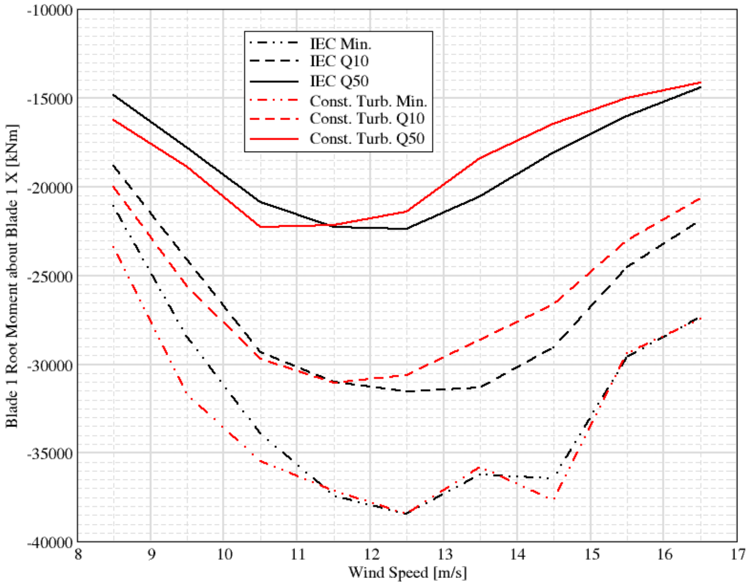
- Tower Top Yaw Moment
  - Below rated, the IEC ECD is larger, above rated the constrained turbulence is larger. These observations hold for both statistical measures.

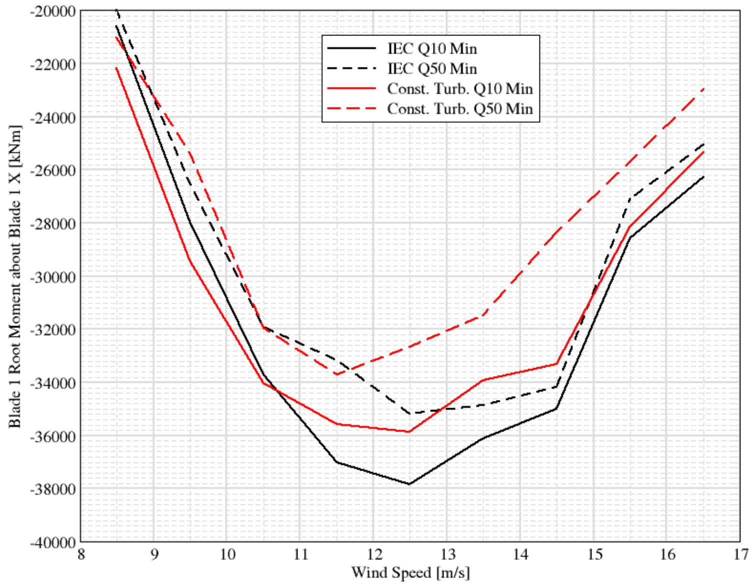


- Shaft Main Bearing Moment XY
  - Less loads overall

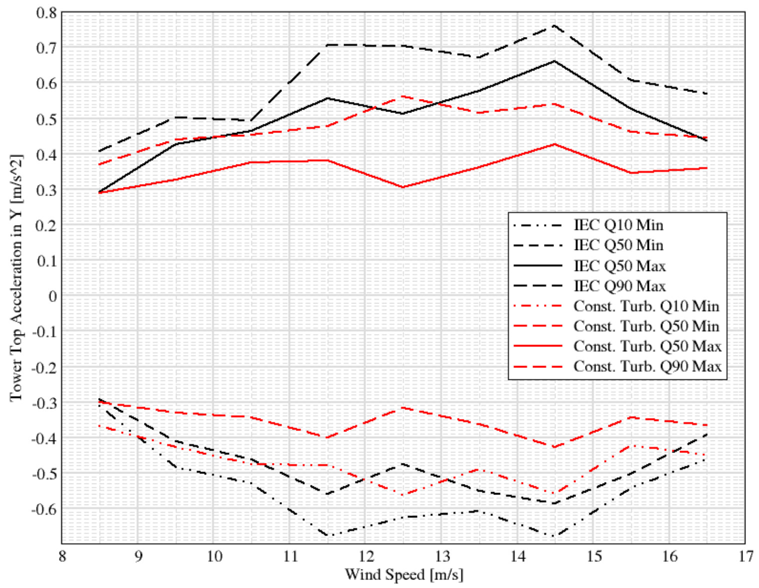


- Blade root flap-wise bending moment
    - ECD appears to be peaking near rated conditions while the constrained turbulence has a broader and flatter peak. Thus, the values tend to be larger for most wind speeds except rated conditions.

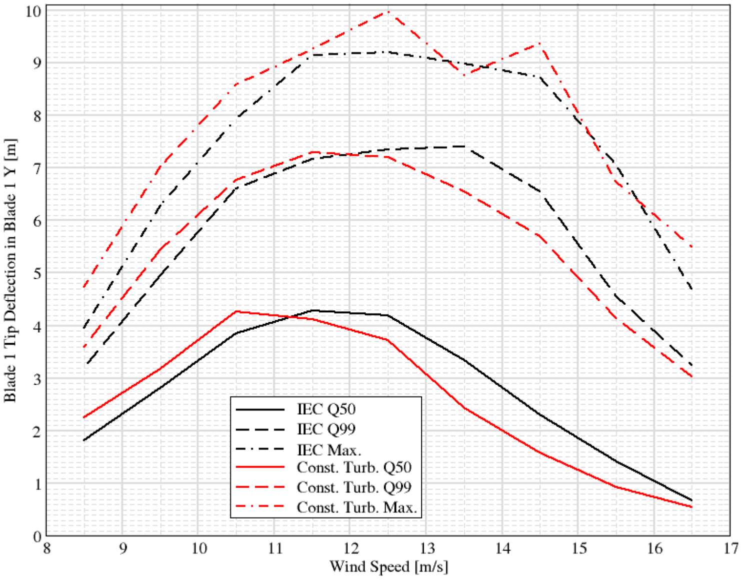


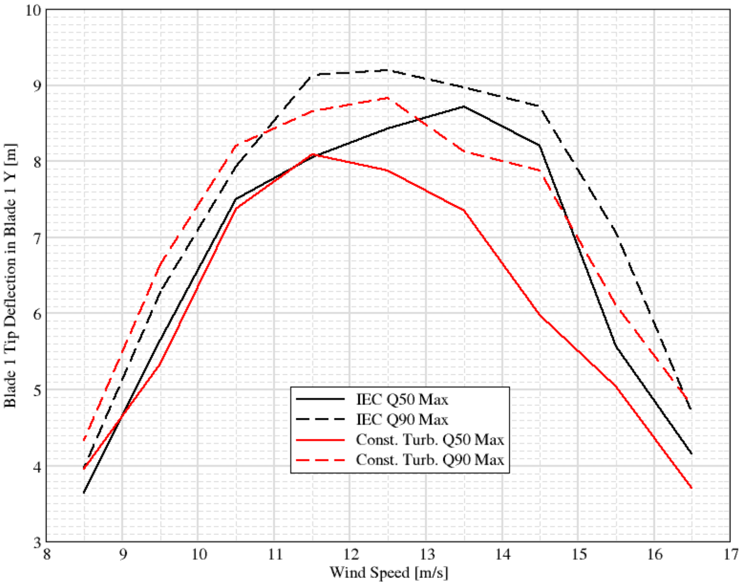


- Tower top acceleration Y
    - Less loads overall



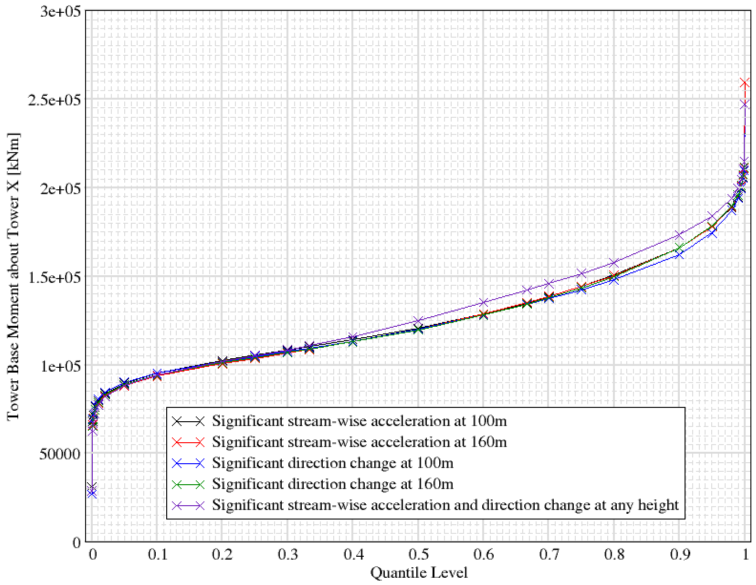
- Tip deflection
    - The conclusion about the loads being less peaked, still holds
    - Below rated, the constrained turbulence loads are similar
    - At rated, the IEC loads are heigher
    - Above rated the constrained loads are higher.



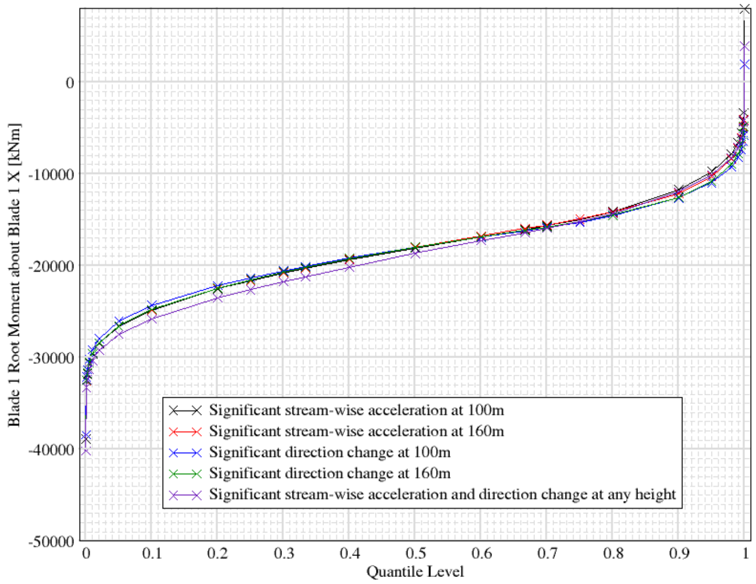


Comparison of different event types for the constrained simulation

- Source is with shear, and the large Mann parameter configuration
- Tower bottom bending moment
    - The auPhi loads appear to be larger
    - The direction change at 100m is the less
    - The remaining loads appear to show similar statistics



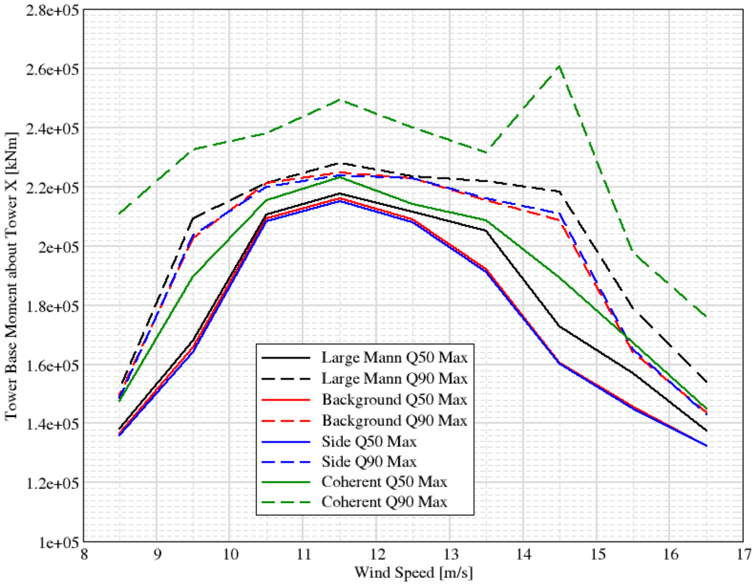
- Root flap-wise bending moment
    - The auPhi loads appear to be larger
    - The direction change at 100m is the less
    - The remaining loads appear to show similar statistics



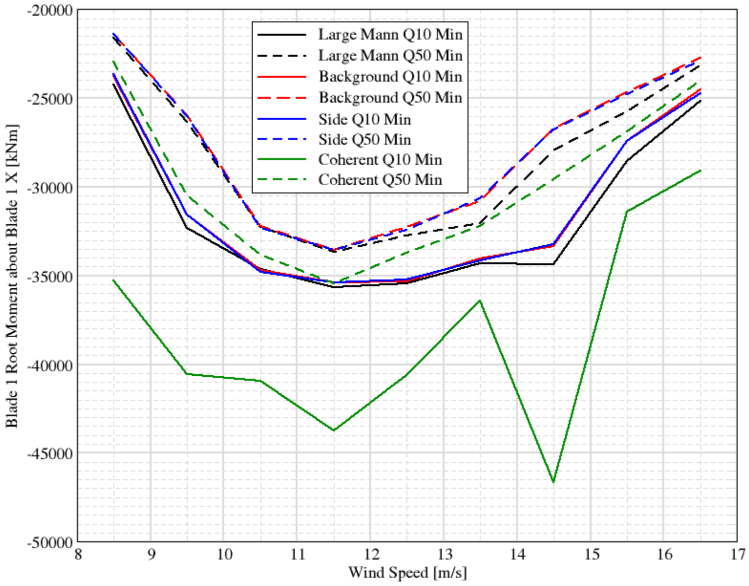
<u>Comparison of different configurations for the constrained simulations</u>

- Source contains all the event types (au100, au160, phi100 and phi160)
- It appears that the coherent simulations may be triggering some shutdowns, furthermore, there were some failed simulations so we can be confident that all simulations were converged
- Tower bottom bending moment

- The maximum of the coherent event maximum appears to be very extreme, so the stats on max appear to be better
- Coherent configuration gives the larger loads
- Then the larger Mann parameter loads, but only for above rated conditions
- Then the side and small Mann parameter events gave smaller and similar loads

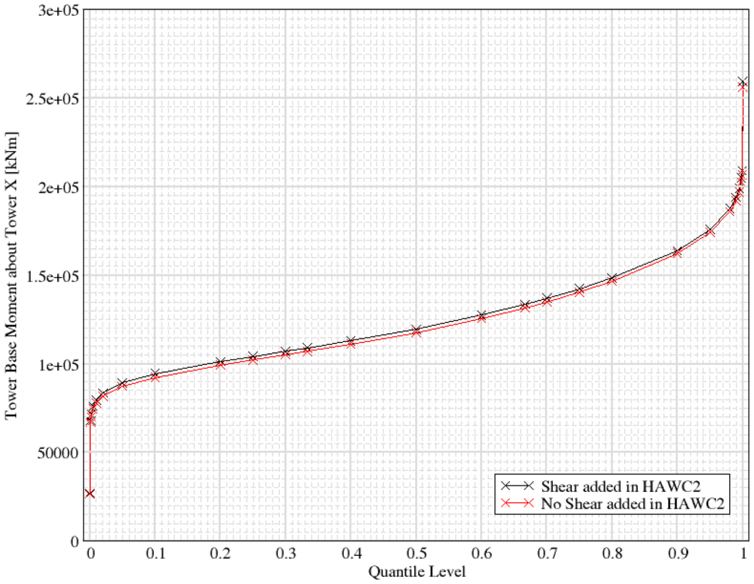


- Blade root flap-wise bending moment
  - The conclusions are similar as with the tower base bending moment



<u>Comparison of the shear vs. no-shear configuration for the constrained simulations</u>

- The source is based on the large Mann configuration and included the loads statistics for all simulations “”
- Tower bottom bending moment
  - It appears that constrained simulations with added shear give slightly larger loads over all quantiles



- Blade root flap-wise bending moment
  - It appears that constrained simulations with added shear give a greater range of loads when additional shear is added.